

From: Davis, Natalie  
Sent: Friday, June 14, 2002 3:29 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/713299

*Please search SEQ ID NO: 1-2 and interference for 09/713299*

Natalie A. Davis, PhD  
Patent Examiner  
Art Unit 1642  
CM1, Rm 8B13  
Mailbox 8E12  
Ph (703) 308-6410

RECEIVED  
JUN 13 2002  
STIC

if Contact:  
Shoppard

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 6/19/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

## TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

## VENDOR/COST(where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:24:13 ; Search time 13.51 Seconds  
(without alignments)  
28.660 Million cell updates/sec

Title: US-09-713-299B-1  
Perfect score: 56  
Sequence: 1 NTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	100.0	10	1	TKS1_AEDAE
2	51	91.1	10	1	TKS2_AEDAE
3	39	69.6	10	1	TKN1_SCYCA
4	38	67.9	845	1	BGLS_KLJMA
5	38	67.9	867	1	SFWD_ECOLI
6	37	66.1	330	1	LAID_PIRAB
7	37	66.1	572	1	VGLF_BRSLA
8	37	66.1	574	1	VGLF_BRSVR
9	36	64.3	11	1	TKNA_RANCA
10	36	64.3	205	1	RS4_RICPR
11	36	64.3	393	1	KGIC_NERDI
12	36	64.3	416	1	SOX_ARATH
13	36	64.3	447	1	YCAJ_ECOLI
14	36	64.3	785	1	CUL3_SCHPO
15	36	64.3	2206	1	POLG_CXA21
16	35	62.5	231	1	YD37_METJA
17	34	60.7	10	1	TKNB_RANCA
18	34	60.7	11	1	TKN_PHYFU
19	34	60.7	12	1	TKN2_KASMA
20	34	60.7	152	1	IPGF_SHIFL
21	34	60.7	152	1	IPGF_SHISO
22	34	60.7	332	1	XYNA_CRYAL
23	34	60.7	381	1	ARGE_BUCAI
24	34	60.7	409	1	PEXC_PICPA
25	34	60.7	482	1	EIBL_ADEL12
26	34	60.7	509	1	ARI2_DROME
27	34	60.7	574	1	VGLF_BRSLA
28	34	60.7	1902	1	P1P_LACLC
29	34	60.7	1902	1	P2P_LACLC
30	34	60.7	1902	1	P2P_LACLC
31	34	60.7	1902	1	P2P_LACLC
32	34	60.7	2214	1	POLG_CXA24
33	33	58.9	14	1	TKNN_RANMA

34	33	58.9	213	1	CAT2_ECOLI	P22615	escherichia
35	33	58.9	270	1	PSFI_DROME	Q9V637	drosophila
36	33	58.9	325	1	LAID_PIRAB	Q57809	pyrococcus
37	33	58.9	332	1	KPPR_SYNV3	P37101	synecocyst
38	33	58.9	357	1	YQO2_CAEEL	Q09305	caenorhabdi
39	33	58.9	375	1	KPPR_CHLRE	P19824	chlamydomon
40	33	58.9	713	1	PMIP_HUMAN	Q99797	homo sapien
41	33	58.9	902	1	UBPF_HUMAN	Q9Y4e8	homo sapien
42	33	58.9	1050	1	LKTA_ACTAC	P16462	actinobacil
43	32	57.1	11	1	TKNA_RANRI	P29207	rana ridibu
44	32	57.1	216	1	FLA3_METJA	Q58303	methanococc
45	32	57.1	255	1	ATP6_YARLI	Q36258	yarrowia li

ALIGNMENTS

RESULT 1	TKS1_AEDAE	STANDARD;	PRT;	10 AA.
AC	P42634;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Sialokinin I.			
OS	Aedes aegypti (Yellowfever mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;			
OC	Aedes.			
OX	NCBI_TaxID=7159;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN=ROCKFELLER; TISSUE=Salivary gland;			
RX	MEDLINE=94105119; PubMed=8278354;			
RA	Champagne D.E., Ribeiro J.M.C.;			
RT	"Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito Aedes aegypti.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).			
CC	-1- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE SITE OF FEEDING.			
CC	-1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.			
DR	PIR; A49581; A49581.			
DR	InterPro: IPR002040; Tachykinin.			
DR	PROSITE: PS00267; TACHYKININ; 1.			
KW	Tachykinin; Neuropeptide; Amidation.			
FT	MOD_RES 10 10 AMIDATION.			
SQ	SEQUENCE 10 AA; 1145 MW; 3DCFDE6B59C33AA8 CRC64;			

Query Match 100.0%; Score 56; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.9e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	NTGDKFYGLM 10		
Db	1	NTGDKFYGLM 10		

RESULT 2	TKS2_AEDAE	STANDARD;	PRT;	10 AA.
AC	P42635;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Sialokinin II.			
OS	Aedes aegypti (Yellowfever mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;			
OC	Aedes.			
OX	NCBI_TaxID=7159;			
RN	[1]			
RP	SEQUENCE.			

RC STRAIN-ROCKEFELLER; TISSUE-Salivary gland;  
RX MEDLINE-94105119; PubMed-8278354;  
RA Champagne D.E., Ribeiro J.M.C.;  
RT "Sialokinin I and II: vasodilatory tachykinins from the yellow fever  
RT mosquito Aedes aegypti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).  
CC -1- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE  
CC SITE OF FEEDING.  
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
DR PIR; B49581; B49581.  
DR InterPro; IPR002040; Tachykinin.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1146 MW; 3DC77C6B59C33AA8 CRC64;

Query Match 91.1%; Score 51; DB 1; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.00037;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
Db 1 DTGDKFYGLM 10  
:|||||

RESULT 3  
TKNL\_SCYCA STANDARD; PRT; 10 AA.  
AC P08608;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Scyllorhinin I.  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;  
OC Scyllorhinidae; Scyllorhinus.  
OX NCBI\_TaxID=7830;  
RN [1]

SEQUENCE.  
RP MEDLINE-86192829; PubMed-2422058;  
RA Conlon J.M., Deacon C.F., O'Toole L., Thim L.;  
RT "Scyllorhinin I and II: two novel tachykinins from dogfish gut.";  
RL FEBS Lett. 200:111-116(1986).  
RN [2]

TISSUE=Brain;  
RC MEDLINE-93292508; PubMed-7685693;  
RA Waugh D., Wang Y., Hazen N., Balmert R.J., Conlon J.M.;  
RT "Primary structures and biological activities of substance-P-related  
RT peptides from the brain of the dogfish, Scyllorhinus canicula.";  
RL Eur. J. Biochem. 214:469-474(1993).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; A24867; A24867.  
DR PIR; S33301; S33301.  
DR InterPro; IPR002040; Tachykinin.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1219 MW; D0602D6B59C33AA9 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
:|||||

Db 4 DKFYGLM 10  
:|||||

RESULT 4  
BGSL\_KLUMA STANDARD; PRT; 845 AA.  
AC P07337;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE Beta-glucosidase precursor (EC 3.2.1.21) (Gentiobiase) (Cellobiase)  
DE (Beta-D-glucoside glucohydrolase).  
OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
OX NCBI\_TaxID=4911;  
RN [1]

SEQUENCE FROM N.A.  
RP STRAIN-ATCC 12424;  
RX MEDLINE-88210533; PubMed-2835179;  
RA Raynal A., Gerbaud C., Francinques M.C., Guerin M.;  
RT "Sequence and transcription of the beta-glucosidase gene of  
RT Kluyveromyces fragilis cloned in Saccharomyces cerevisiae.";  
RL Curr. Genet. 12:175-184(1987).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-  
CC glucose residues with release of beta-D-glucose.  
CC -1- PATHWAY: CELLULOSE DEGRADATION.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X05918; CAA29353.1; -  
DR PIR; A29148; GLVK.  
DR InterPro; IPR001764; Glyco\_hydro\_3.  
DR InterPro; IPR002772; Glyco\_hydro\_3C.  
DR Pfam; PF00933; Glyco\_hydro\_3; 1.  
DR Pfam; PF01915; Glyco\_hydro\_3\_C; 1.  
DR PRINTS; PR00133; GLHYDRLASE3.  
DR PROSITE; PS00775; GLYCOSYL\_HYDROL\_F3; 1.  
DR Hydrolase; Glycosidase; Cellulose degradation; Glycoprotein; Signal.  
FT SIGNAL 1 ?  
FT CHAIN ? 845  
FT ACT\_SITE 225 225 BY SIMILARITY.  
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 845 AA; 93916 MW; 78D4C15D05992B5 CRC64;

Query Match 67.9%; Score 38; DB 1; Length 845;  
Best Local Similarity 87.5%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8  
:|||||  
Db 745 NTGDKFAG 752

RESULT 5  
SFMD\_ECOLI STANDARD; PRT; 867 AA.  
ID SFMD\_ECOLI  
AC P77468; P77133;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)

OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ORSAY;  
RA Heilig R.;  
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
structure and evolution.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RR 1- CATALYTIC ACTIVITY: L-aminocyclopropane-L-carboxylate + H(2)O = 2-oxobutanolate + NH(3).  
CC -| COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC -| SIMILARITY: BELONGS TO THE ACC DEAMINASE FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
CC EMBL: AJ248283; CAB48986.1; .  
DR InterPro: IPR001926; PALP.  
DR Pfam: PF00291; PALP; 1.  
KW Hypothetical protein; Lyase; Pyridoxal phosphate; Complete proteome.  
FT BINDING 54 54 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 330 AA; 35754 MW; CC8699923C0B11CD CRC64;  
  
Query Match 66.1%; Score 37; DB 1; Length 330;  
Best Local Similarity 77.8%; Pred. No. 7.3;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 TGDKFYGLM 10  
|| |||||  
Db 287 TGAFUYGLM 295  
  
RESULT 7  
VGLF\_BRSVA STANDARD; PRT; 572 AA.  
AC E29791;  
AD 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
Fusion glycoprotein Fl].  
DE F.  
OS Bovine respiratory syncytial virus (strain A51908) (BRs).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11247;  
XN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=92185490; PubMed=1312130;  
RA Zamora M., Samal S.K.;  
RW "Sequence analysis of M2 mRNA of bovine respiratory syncytial virus  
obtained from an F-M2 dicistronic mRNA suggests structural homology  
with that of human respiratory syncytial virus.";  
RL J. Gen. Virol. 73:737-741(1992).  
CC -| FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
MEMBRANES, RESULTING IN VIRAL PENETRATION, & CAN DIRECT FUSION OF  
INFECTED CELLS WITH ADJOINING CELLS, RESULTING IN THE FORMATION OF  
SYNCYTIUM.  
CC -| SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
LINKED BY A DISULFIDE BOND.  
CC -| SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M82816; AAA42804.1; -;  
DR PIR; JQ1481; VGNZBA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; fusion\_gly; 1.  
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal;  
KW Lipoprotein; Palmitate.  
FT SIGNAL 1 25  
FT CHAIN 26 572  
FT CHAIN 26 136  
FT CHAIN 137 572  
FT CHAIN 137 158  
FT TRANSMEM 520 547  
FT DOMAIN 550 572  
FT LIPID 548 548  
FT CARBOHYD 27 27  
FT CARBOHYD 70 70  
FT CARBOHYD 120 120  
FT CARBOHYD 498 498  
SQ SEQUENCE 572 AA; 63443 MW; A055E5E65801663E CRC64;

Query Match 66.1%; Score 37; DB 1; Length 572;  
Best Local Similarity 70.0%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NTGDKFYGLM 10  
Db 120 NSTKKFYGLM 129

RESULT 8  
VGLF\_BRSVR  
ID VGLF\_BRSVR STANDARD; PRT; 574 AA.  
AC P23728;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN F.  
OS Bovine respiratory syncytial virus (strain B94) (BRS).  
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11249;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91108398; PubMed=1703214;  
RA Walravens K., Kettmann R., Collard A., Coppe P., Burny A.;  
RT "Sequence comparison between the fusion protein of human and bovine  
RT respiratory syncytial viruses.";  
RL J. Gen. Virol. 71:3009-3014(1990).  
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
CC MEMBRANES, RESULTING IN VIRAL PENETRATION, & CAN DIRECT FUSION OF  
CC INFECTED CELLS WITH ADJOINING CELLS, RESULTING IN THE FORMATION OF  
CC SYNCYTIA.  
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
CC LINKED BY A DISULFIDE BOND.  
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
CC FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR PIR; JH0233; VGNZBS.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; fusion\_gly; 1.  
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal;  
KW Lipoprotein; Palmitate.  
FT SIGNAL 1 25  
FT CHAIN 26 574  
FT CHAIN 26 136  
FT CHAIN 137 574  
FT TRANSMEM 137 158  
FT TRANSMEM 522 549  
FT DOMAIN 550 574  
FT LIPID 550 550  
FT CARBOHYD 27 27  
FT CARBOHYD 70 70  
FT CARBOHYD 120 120  
FT CARBOHYD 500 500  
SQ SEQUENCE 574 AA; 63630 MW; 856718D3E7AE04AE CRC64;

Query Match 66.1%; Score 37; DB 1; Length 574;  
Best Local Similarity 70.0%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NTGDKFYGLM 10  
Db 120 NSTKKFYGLM 129

RESULT 9  
TKNA\_RANCA  
ID TKNA\_RANCA STANDARD; PRT; 11 AA.  
AC P22688;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Ranatachykinin A (RTK A).  
OS Rana catesbeiana (Bull frog).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranolia; Ranidae; Rana.  
OX NCBI\_TaxID=8400;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Brain, and Intestine;  
RX MEDLINE=91254337; PubMed=2043143;  
RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;  
RT "Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and  
RT brain and intestine.";  
RL Biochem. Biophys. Res. Commun. 177:588-595(1991).  
RN [2]  
RP SEQUENCE.  
RC TISSUE=Intestine;  
RX MEDLINE=94023216; PubMed=8210506;  
RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;  
RT "Four novel tachykinins in frog (Rana catesbeiana) brain and  
RT intestine.";  
RL Regul. Pept. 46:81-88(1993).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; JQ0426; JE0426.  
DR PIR; A61033; A61033.  
DR InterPro; IPR003580; Protachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 11 11  
SQ SEQUENCE 11 AA; 1311 MW; 200D60CC59D40AB7 CRC64;

```

Query Match      64.3%; Score 36; DB 1; Length 11;
Best Local Similarity 85.7%; Pred. No. 0.35;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 4 DKFYGLM 10
Db 5 DRFYGLM 11

RESULT 10
RS4_RICPR STANDARD; PRT; 205 AA.
AC Q9ZDI3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S4.
GN RPSD OR RP345.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sierichitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA
CC (BY SIMILARITY)
CC -!- SIMILARITY: CONTAINS 1 S4 RNA-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ235271; CAA14805.1; -.
CC HSP; P81288; IC05.
CC InterPro; IPR001912; Ribosomal_S4.
CC InterPro; IPR002942; S4.
CC Pfam; PF00163; Ribosomal_S4; 1.
CC Pfam; PF01479; S4; 1.
CC SMART; SM00363; S4; 1.
CC PROSITE; PS00632; RIBOSOMAL_S4; 1.
CC Ribosomal protein; rRNA-binding; Complete proteome.
CC DOMAIN 94 RNA-BINDING (S4 TYPE).
CC SEQUENCE 205 AA; 23318 MW; 3750C8D28A9FAF49 CRC64;

Query Match      64.3%; Score 36; DB 1; Length 205;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
Db 83 NTGENFIGLL 92

RESULT 11
KGCY_NERDI STANDARD; PRT; 393 AA.
AC P51546;
DT 01-OCT-1996 (Rel. 34, Created)

```

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycocyamine kinase (EC 2.7.3.1) (GK) (Guanidoacetate kinase).
OS Nereis diversicolor (Sandworm) (Hediste diversicolor).
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC Phyllocorida; Nereididae; Neanthes.
OX NCBI_TaxID=6352;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94194511; PubMed=8145248;
RA Suzuki T., Furukohri T.;
RT "Evolution of phosphagen kinase. Primary structure of glycocyamine
RT kinase and arginine kinase from invertebrates."
RL J. Mol. Biol. 237:353-357(1994).
CC -!- CATALYTIC ACTIVITY: ATP + guanidoacetate = ADP +
CC phosphoguanidoacetate.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE ATP:GUANIDO PHOSPHOTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D26103; BAA05099.1; -.
CC HSP; P11009; ICRK.
CC InterPro; IPR000749; ATP-gua_Ptrans.
CC Pfam; PF00217; ATP-gua_Ptrans; 1.
CC Pfam; PF02807; ATP-gua_Ptrans; 1.
CC PROSITE; PS00112; GUANIDO_KINASE; 1.
CC Transferase; Kinase.
KW ACT_SITE 278 278 BY SIMILARITY
FT ACT_SITE 393 AA; 44461 MW; CB529ADC9B3CCFC1 CRC64;
SQ SEQUENCE 393 AA; 44461 MW; CB529ADC9B3CCFC1 CRC64;

Query Match      64.3%; Score 36; DB 1; Length 393;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8
Db 58 NPGNKFYG 65

RESULT 12
SOX_ARATH STANDARD; PRT; 416 AA.
AC Q9SJA7;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Potential sarcosine oxidase (EC 1.5.3.1).
GN AT2G24580 OR F25P17.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;

```

```

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RT Nature 402:761-768(1999).
CC -I- CATALYTIC ACTIVITY: Sarcosine + H(2)O + O(2) = glycine +
CC formaldehyde + H(2)O(2).
CC -I- COFACTOR: FAD; 1 mole of FAD per mole of enzyme (By similarity).
CC -I- SIMILARITY: BELONGS TO THE MSOX/MTOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AC006954; AAD23888.1; -.
CC DR HSSP; P40859; 1B3M.
CC DR InterPro; IPR000205; NAD_binding.
KW Hypothetical protein; Oxidoreductase; FAD; Flavoprotein.
FT NP_BIND 10 40 FAD (ADP PART) (POTENTIAL).
FT NP_BIND 325 325 FAD (COVALENT) (PROBABLE).
SQ SEQUENCE 416 AA; 45702 MW; E12B3792DF305BCC CRC64;

```

Query Match 64.3%; Score 36; DB 1; Length 416;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 GDKFYG 8
Db 197 GDKFYG 202

```

RESULT 13

ID	YCAJ_ECOLI	STANDARD;	PRT;	447 AA.
AC	P45936; P75933;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein ycaj			
GN	YCAJ OR B0892 OR 21238 OR ECS0977.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RT	Science 277:1453-1474(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=97061202; PubMed=8905232;			
RA	Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,			
RA	Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,			
RA	Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,			
RA	Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,			
RA	Sampe G., Seki Y., Tagami H., Takenoto K., Wada C., Yamamoto Y.,			
RA	Yano M., Horiuchi T.;			
RT	"A 718-kb DNA sequence of the Escherichia coli K-12 genome			
RT	corresponding to the 12.7-28.0 min region on the linkage map.";			
RL	DNA Res. 3:137-155(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A.			

```

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

```

"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
Nature 409:529-533(2001).  
[4]  
SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;

"Complete genome sequence of enterohaemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12.";  
DNA Res. 8:11-22(2001).  
[5]  
SEQUENCE OF 1-61 FROM N.A.  
RC STRAIN=K12 / MC4100;  
RX MEDLINE=95354654; PubMed=7628437;  
RA Matsuyama S.-I., Tajima T., Tokuda H.;

"A novel periplasmic carrier protein involved in the sorting and  
transport of Escherichia coli lipoproteins destined for the outer  
membrane.";  
EMBO J. 14:3365-3372(1995).  
[6]  
SEQUENCE OF 367-447 FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=87146412; PubMed=3029694;  
RA Haertlein M., Madern D., Leberman R.;

"Cloning and characterization of the gene for Escherichia coli seryl-  
tRNA synthetase.";  
Nucleic Acids Res. 15:1005-1017(1987).  
[7]  
IDENTIFICATION.  
RX MEDLINE=96032851; PubMed=7567469;  
RA Borodovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,  
RA Danchin A.;

"Detection of new genes in a bacterial genome using Markov models for  
three gene classes.";  
Nucleic Acids Res. 23:3554-3562(1995).  
RT -I- SIMILARITY: STRONG TO H INFLUENZA H11590 AND C. BURNETII HOMOLOG.  
CC -I- SIMILARITY: TO YEAST YNL218W.  
CC -I- CAUTION: REF.6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT IN POSITION 370.  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

EMBL; AE000191; AAC73978.1; -.
EMBL; D90726; BAA35617.1; -.
EMBL; D90727; BAA35624.1; -.
EMBL; AE005278; AAG55379.1; -.
EMBL; AP002553; BAB34400.1; -.
EMBL; D49398; -. NOT_ANNOTATED_CDS.
EMBL; X05017; -. NOT_ANNOTATED_CDS.
EcGene; EG12690; ycaJ.
InterPro; IPR003593; AAA.
InterPro; IPR003959; AAA_subfam.
InterPro; IPR000862; REC.
Pfam; PF00004; AAA; 1.
SMART; SM00382; AAA; 1.
Hypothetical protein; ATP-binding; Complete proteome.

FT NP\_BIND 57 64 ATP (POTENTIAL);  
SQ SEQUENCE 447 AA; 49626 MW; 116ACBE38E4E4A3A CRC64;

Query Match 64.3%; Score 36; DB 1; Length 447;  
Best Local Similarity 60.0%; Pred. No. 16;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
| | | | |  
Db 250 NKGDRFYDLI 259

## RESULT 14

CUL3\_SCHPO .  
ID CUL3\_SCHPO STANDARD; PRT; 785 AA.  
AC Q09760; O74185;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cullin homolog 3 (CUL-3).  
GN PCU3 OR SPAC24H6.03.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

RN [2]  
RP SEQUENCE OF 136-543 FROM N.A.  
RA Koninani K., Toda T.;

RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CULLIN FAMILY.  
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 513  
CC ONWARD AND IS SHORTER (421 AA) DUE TO A FRAMESHIFT.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; 254142; CAA90847.1; .

DR EMBL; AB017028; BAA32519.1; ALT\_FRAME.

DR InterPro; IPR001373; Cullin.

DR Pfam; PF00888; Cullin; 1.

DR SMART; SM00182; CULLIN; 1.

DR PROSITE; PS01256; CULLIN\_1; 1.

DR PROSITE; PS00069; CULLIN\_2; 2.

DR Hypothetical protein.

FT CONFLICT 476 476 R -> RYALIVETVENTER (IN REF. 2).

SQ SEQUENCE 785 AA; 90511 MW; 0830F7A857DF6CE7 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 785;  
Best Local Similarity 85.7%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFY 7  
| | | | |

Db 671 NAGDKFY 677

## RESULT 15

POLG\_CXA21  
ID POLG\_CXA21 STANDARD; PRT; 2206 AA.  
AC P22055;

DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2  
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
OS Coxsackievirus A21 (strain Coe).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=12070;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90063544; PubMed=2584950;  
RA Hughes P.J., North C., Minor P.D., Stanway G.;  
RT "The complete nucleotide sequence of coxsackievirus A21.";  
RL J. Gen. Virol. 70:2943-2952(1989).  
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR  
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.  
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; D00538; BAA00426.1; .

DR PIR; A33373; GNNY21.

DR HSPP; P21404; ID4M.

DR MEROPS; C03.011; .

DR MEROPS; C03.022; .

DR InterPro; IPR000199; Cys-protease-3C.

DR InterPro; IPR00138; Pico\_P1A.

DR InterPro; IPR000081; Pico\_P2A.

DR InterPro; IPR002527; Pico\_P2B.

DR InterPro; IPR000605; RNA\_helicase.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR001676; Rhv.

DR Pfam; PF00548; Cys-protease-3C; 1.

DR Pfam; PF02226; Pico\_P1A; 1.

DR Pfam; PF00947; Pico\_P2A; 1.

DR Pfam; PF01532; Pico\_P2B; 1.

DR Pfam; PF00073; rhv; 3.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

DR ProDom; PD001125; Cys-protease-3C; 1.

DR ProDom; PD001274; Pico\_P2B; 1.

DR ProDom; PD001306; Pico\_P2A; 1.

KW Polyprotein; Coat protein; Core protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.

FT CHAIN 2 69

FT CHAIN 70 341

FT CHAIN 342 581

FT CHAIN 582 879

FT CHAIN 880 1028

FT CHAIN 1029 1125

FT CHAIN 1126 1453

FT CHAIN 1454 1540

FT CHAIN 1541 1562

FT CHAIN 1563 1745

FT CHAIN 1746 2206

FT LIPID 2 2

FT ACT\_SITE 1709 1709



FT ACT\_SITE 1723 1723 PROTEASE (POTENTIAL).  
SQ SEQUENCE 2206 AA; 246049 MW; 537F6A13AD37723B CRC64;

Query Match 64.3%; Score 36; DB 1; Length 2206;  
Best Local Similarity 60.0%; Pred. No. 80;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
: || || || :  
Db 799 DAGDTFYGLV 808

Search completed: June 19, 2002, 08:29:55  
Job time: 342 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:23:13 ; Search time 41.44 Seconds  
(without alignments)  
41.746 Million cell updates/sec

Title: US-09-713-299B-1  
Perfect score: 56  
Sequence: 1 NTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	56	100.0	85	5	Q9TW63	Q9tw63	aedes aegyp
2	56	100.0	85	5	Q9UAR7	Q9uar7	aedes aegyp
3	51	91.1	11	5	Q9UAR8	Q9uar8	aedes aegyp
4	42	75.0	1157	16	Q97F62	Q97f62	clostridium
5	42	75.0	1217	16	Q97GW3	Q97gw3	clostridium
6	41	73.2	697	13	Q98TV9	Q98tv9	notothenia
7	41	73.2	697	13	Q98TV4	Q98tv4	chaenoceph
8	41	73.2	697	13	Q98TV3	Q98tv3	gobionototh
9	40	71.4	405	10	Q9ATC3	Q9atc3	vaucheria 1
10	38	67.9	224	2	P77133	P77133	escherichia
11	38	67.9	415	16	Q97IV5	Q97iv5	clostridium
12	38	67.9	509	8	Q95ED5	Q95ed5	blossfeldia
13	38	67.9	509	8	Q95ED4	Q95ed4	blossfeldia
14	37	66.1	304	12	Q9YLN6	Q9yln6	human coxa
15	37	66.1	305	12	Q9YLM6	Q9ylm6	human coxa
16	37	66.1	305	12	Q9YLJ4	Q9ylj4	echovirus 3

17	37	66.1	361	12	Q9QAA9	Q9qaa9	bovine resp
18	37	66.1	435	10	Q942T2	Q942t2	oryza sativ
19	37	66.1	440	5	O96686	O96686	oryza sativ
20	37	66.1	440	5	O9VJ02	O9vj02	oryza sativ
21	37	66.1	478	16	O9JZ64	O9jz64	neisseria m
22	37	66.1	478	16	O9JU61	O9ju61	neisseria m
23	37	66.1	539	16	O9HXT2	O9hxt2	pseudomonas
24	37	66.1	574	12	O9O698	O9o698	bovine resp
25	37	66.1	574	12	O9XS24	O9xs24	bovine resp
26	37	66.1	1134	17	O30078	O30078	archaeoglob
27	36	64.3	92	13	O98TV8	O98tv8	ginglymosto
28	36	64.3	138	12	O9JA93	O9ja93	human coxa
29	36	64.3	139	12	O9W902	O9w902	human coxa
30	36	64.3	140	12	O9JAA0	O9jaa0	human coxa
31	36	64.3	140	12	O9JA99	O9ja99	human coxa
32	36	64.3	140	12	O9JA58	O9ja58	human coxa
33	36	64.3	205	16	O92IF2	O92if2	rickettsia
34	36	64.3	244	12	O994Y5	O994y5	sacbrood vi
35	36	64.3	244	12	O994Y4	O994y4	sacbrood vi
36	36	64.3	244	12	O994Y3	O994y3	sacbrood vi
37	36	64.3	244	12	O994Y2	O994y2	sacbrood vi
38	36	64.3	244	12	O994Y1	O994y1	sacbrood vi
39	36	64.3	244	12	O994Y0	O994y0	sacbrood vi
40	36	64.3	244	12	O994X9	O994x9	sacbrood vi
41	36	64.3	244	12	O994X8	O994x8	sacbrood vi
42	36	64.3	244	12	O994X7	O994x7	sacbrood vi
43	36	64.3	244	12	O994X5	O994x5	sacbrood vi
44	36	64.3	244	12	O994X4	O994x4	sacbrood vi
45	36	64.3	244	12	O994X3	O994x3	sacbrood vi

ALIGNMENTS

RESULT 1

ID Q9TW63 PRELIMINARY; PRT; 85 AA.  
AC Q9TW63;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)  
DE SIALOKININ I PREPROTEIN.  
OS Aedes aegypti (Yellow fever mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;  
OC Aedes.  
OX NCBI\_TaxID=7159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ROCKEFELLER;  
RA Beerntsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;  
RT "Characterization of the sialokinin I gene encoding the salivary  
RT vasodilator of the yellow fever mosquito, Aedes aegypti.";  
RL Insect Mol. Biol. 0:0-0(1999).  
DR EMBL; AF108102; AAD16886.1; -  
DR EMBL; AF108101; AAD16885.1; -  
DR InterPro; IPR002040; Tachykinin.  
DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_1.  
SQ SEQUENCE 85 AA; 9326 MW; 209F03E3932A6A8B CRC64;

Query Match 100.0%; Score 56; DB 5; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
|||||||  
Db . 75 NTGDKFYGLM 84

RESULT 2

Q9UAR7 PRELIMINARY; PRT; 85 AA.  
ID Q9UAR7

```
AC Q9UAR7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SALIVARY VASODILATORY PROTEIN PRECURSOR.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ROCKEFELLER;
RX MEDLINE=20099025; PubMed=10620041;
RA Beertsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT "Characterization of the Sialokinin I gene encoding the salivary
RT vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL Insect Mol. Biol. 8:459-467(1999).
DR EMBL; AF108099; AAD17916.1; -.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 75 85 SALIVARY VASODILATORY PROTEIN.
FT SEQUENCE 85 AA; 9460 MW; EDA669E3470638DF CRC64;

Query Match 100.0%; Score 56; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
DB 75 :NTGDKFYGLM 84

RESULT 3
Q9UAR8 PRELIMINARY; PRT; 11 AA.
AC Q9UAR8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE STALOKININ I PREPROTEIN (FRAGMENT).
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ROCKEFELLER/RED; TISSUE=SALIVARY GLAND;
RX MEDLINE=20099025; PubMed=10620041;
RA Beertsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT "Characterization of the Sialokinin I gene encoding the salivary
RT vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL Insect Mol. Biol. 8:459-467(1999).
DR EMBL; AF108100; AAD16884.1; -.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
FT NON_TER 1 1
FT SEQUENCE 11 AA; 1203 MW; 8BAD77C6B59C33A CRC64;
```

```
Query Match 91.1%; Score 51; DB 5; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.0026;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
DB 1 DTGDKFYGLM 10
```

```
RESULT 4
Q97F52 PRELIMINARY; PRT; 1157 AA.
AC Q97F62;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE FUSION OF ALPHA-GLUCOSIDASE (FAMILY 31 GLYCOSYL HYDROLASE) AND
DE GLYCOSIDASE (TREA/MALS FAMILY).
GN CAC2891.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R.L., Sabathe F., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007786; AAK80833.1; -.
DR InterPro; IPR000322; Glyco_hydro_31.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF01055; Glyco_hydro_31; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 1157 AA; 129411 MW; 1DB1728D833F7578 CRC64;
```

```
Query Match 75.0%; Score 42; DB 16; Length 1157;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 NTGDKFYGL 9
DB 165 NSGDRFYGI 173

RESULT 5
Q97GW3 PRELIMINARY; PRT; 1217 AA.
AC Q97GW3;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALPHA-GLUCOSIDASE FUSED TO UNKNOWN ALPHA-AMYLASE C-TERMINAL.
DE DOMAIN.
GN CAC2252.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007725; AAK80209.1; -.
DR InterPro; IPR000322; Glyco_hydro_31.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF01055; Glyco_hydro_31; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
```

KW Complete proteome.  
SQ SEQUENCE 1217 AA; 134530 MW; 26F9D66B02688A30 CRC64;

Query Match 75.0%; Score 42; DB 16; Length 1217;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTGDKFYGL 9  
|:|:|:|:|:  
Db 169 NSGDREYGI 177

RESULT 6  
Q98TV9 PRELIMINARY; PRT; 697 AA.  
ID Q98TV9  
AC Q98TV9  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE LONG CHAIN FATTY ACYL COA SYNTHETASE (EC 6.2.1.3).  
OS Notothenia coriiceps (black rockcod).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;  
OC Notothenioidel; Nototheniidae; Notothenia.  
OX NCBI\_TaxID=8208;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PECTORAL ADDUCTOR PROFUNDUS;  
RA Grove T.J., Sidell B.D.;  
RT "Fatty Acyl CoA Synthetase in Antarctic Fishes";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF326574; AAG49599.1; -;  
DR HSSP; P08659; ILCI.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PRINTS; PR00154; AMPBINDING.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Ligase.  
SQ SEQUENCE 697 AA; 77807 MW; E5B273B706F7DF06 CRC64;

Query Match 73.2%; Score 41; DB 13; Length 697;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NTGDKFYGL 9  
|:|:|:|:|:  
Db 144 NTGDKFVGI 152

RESULT 7  
Q98TV4 PRELIMINARY; PRT; 697 AA.  
ID Q98TV4  
AC Q98TV4  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE LONG CHAIN FATTY ACYL COA SYNTHETASE (EC 6.2.1.3).  
OS Chaenocephalus aceratus (White crocodile fish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;  
OC Notothenioidel; Channichthyidae; Chaenocephalus.  
OX NCBI\_TaxID=36190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PECTORAL ADDUCTOR PROFUNDUS;  
RA Grove T.J., Sidell B.D.;  
RT "Fatty Acyl CoA Synthetase in Antarctic Fishes";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF327744; AAK07470.1; -;

DR HSSP; P08659; ILCI.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PRINTS; PR00154; AMPBINDING.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Ligase.  
SQ SEQUENCE 697 AA; 77766 MW; FA77D01967DC3E57 CRC64;

Query Match 73.2%; Score 41; DB 13; Length 697;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NTGDKFYGL 9  
|:|:|:|:|:  
Db 144 NTGDKFICI 152

RESULT 8  
Q98TV3 PRELIMINARY; PRT; 697 AA.  
ID Q98TV3  
AC Q98TV3  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE LONG CHAIN FATTY ACYL COA SYNTHETASE (EC 6.2.1.3).  
OS Gobionotothen gibberifrons (Humped rockcod).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;  
OC Notothenioidel; Nototheniidae; Gobionotothen.  
OX NCBI\_TaxID=36202;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PECTORAL ADDUCTOR PROFUNDUS;  
RA Grove T.J., Sidell B.D.;  
RT "Fatty Acyl CoA Synthetase in Antarctic Fishes";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF327745; AAK07471.1; -;  
DR HSSP; P08659; ILCI.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PRINTS; PR00154; AMPBINDING.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Ligase.  
SQ SEQUENCE 697 AA; 77607 MW; BE85DCED5A687695 CRC64;

Query Match 73.2%; Score 41; DB 13; Length 697;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NTGDKFYGL 9  
|:|:|:|:|:  
Db 144 NTGDKFICI 152

RESULT 9  
Q9ATC3 PRELIMINARY; PRT; 405 AA.  
ID Q9ATC3  
AC Q9ATC3  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE PHOSPHORIBULOKINASE.  
OS Vaucheria litorea.  
OC Eukaryota; stramenopiles; Xanthophyceae; Vaucheriales; Vaucheriaceae;  
OC Vaucheria.  
OX NCBI\_TaxID=109269;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Summer E.J., Rumpho M.E.;  
RT "Chloroplast localized, nuclear encoded proteins persist for many

RT months in an animal cell despite the lack of cognate algal nuclear  
 genes.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBDJ databases.  
 DR EMBL: AF336986; AAK21910.1; -  
 DR InterPro: IPR001324; PRK;  
 DR InterPro: IPR000764; Uridine\_kin.  
 DR Pfam: PF00485; PRK; 1.  
 DR PRINTS; PR00478; PHRIDINKINASE.  
 DR PRINTS; PR00988; URIDINKINASE.  
 DR PROSITE; PS00567; PHOSPHORIBULOKINASE; 1.  
 KW Kinase.  
 SQ SEQUENCE 405 AA; 44805 MW; 0AED85D4071142BF CRC64;

Query Match 71.4%; Score 40; DB 10; Length 405;  
 Best Local Similarity 87.5%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8

|||||

Db 353 NTGTFKYG 360

RESULT 10

ID P77133 PRELIMINARY; PRT; 224 AA.

AC P77133;

DT 01-FEB-1997 (TREMBlrel. 02, Created)

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE SIMILAR TO S. TYPHIMURIUM FIMD (FRAGMENT).

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,

RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,

RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;

RT "Sequence of minutes 4-25 of Escherichia coli."

RL Submitted (DEC-1996) to the EMBL/GenBank/DBDJ databases.

DR EMBL: U82664; AAB40285.1; -

DR InterPro: IPR000015; Fimb\_usher.

DR Pfam: PF00577; Usher; 1.

FT NON\_TER 224 224

SQ SEQUENCE 224 AA; 25068 MW; FE66069677EADE27 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 224;  
 Best Local Similarity 80.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10

|||||

Db 92 NTGDKSGGLM 101

RESULT 11

ID Q971Y5 PRELIMINARY; PRT; 415 AA.

AC Q971Y5;

DT 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE DNA-METHYLTRANSFERASE (CYTOSINE-SPECIFIC), ORTHOLOG OF BSP6I

DE BSBUTILIS.

GN CAC1501.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

DE

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466286;  
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 bacterium Clostridium acetobutylicum.";  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE007660; AAK79469.1; -  
 DR InterPro: IPR001525; C5\_DNA\_meth.  
 DR Pfam: PF00145; DNA\_methylase; 2.  
 DR PRINTS; PR00105; C5METHYTRFRASE.  
 DR PROSITE; PS00094; C5\_MTASE\_1; 1.  
 DR PROSITE; PS00095; C5\_MTASE\_2; UNKNOWN\_1.  
 KW Transferase; Methyltransferase; Complete proteome.  
 SQ SEQUENCE 415 AA; 48320 MW; C6D029B9ED3901F3 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 415;  
 Best Local Similarity 85.7%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFY 7

|||||

Db 266 NTGDKYY 272

RESULT 12

ID Q95ED5 PRELIMINARY; PRT; 509 AA.

AC Q95ED5;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE MATURASE K.

GN MATK.

OS Blossfeldia liliputana.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllidae; Caryophyllales; Cactaceae; Blossfeldia.

OX NCBI\_TaxID=154411;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BLOLILE215;

RA Nyffeler R.;

RT "Phylogenetic relationships in the cactus family (Cactaceae) based on  
 evidence from trnK/matK and trnL-trnF sequences.";  
 RL Am. J. Bot. 0:0-0(2001).  
 DR EMBL: AY015283; AAK19770.1; -  
 KW Chloroplast.

SQ SEQUENCE 509 AA; 60820 MW; 1459F764196122FA CRC64;

Query Match 67.9%; Score 38; DB 8; Length 509;  
 Best Local Similarity 70.0%; Pred. No. 52;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10

|||||

Db 45 NTGDKKYSLL 54

RESULT 13

ID Q95ED4 PRELIMINARY; PRT; 509 AA.

AC Q95ED4;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE MATURASE K.

```
GN MATK.
OS Blossfeldia liliuputana.
OG Chloroplast.
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Cactaceae; Blossfeldia.
OX NCBI_TaxID=15441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL01LE209;
RA Nyfeler R.;
RT "Phylogenetic relationships in the cactus family (Cactaceae) based on
RT evidence from trnK/matK and trnL-trnF sequences.";
RL Am. J. Bot. 0:0-0(2001).
DR EMBL; AY015284; AAK19771.1; -.
KW Chloroplast.
SQ SEQUENCE 509 AA; 60770 MW; 93EB1061487D94E9 CRC64;

Query Match 67.9%; Score 38; DB 8; Length 509;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
Db 45 NTGDKKYSLL 54

RESULT 14
QYLN6 PRELIMINARY; PRT; 304 AA.
AC Q9YLN6;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A11.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=42770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BELGIUM-1;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081301; RAD17705.1; -.
DR HSSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 304 304
SQ SEQUENCE 304 AA; 33884 MW; F151A7C24CD7A4C1 CRC64;

Query Match 66.1%; Score 37; DB 12; Length 304;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYGL 9
Db 225 DTGDAYYGL 233

RESULT 15
QYLM6 PRELIMINARY; PRT; 305 AA.
ID Q9YLM6;
AC Q9YLM6;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
```

```
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A24.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12089;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JOSEPH;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081311; AAD17715.1; -.
DR HSSP; P03300; 2PLV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 305 305
SQ SEQUENCE 305 AA; 34091 MW; D4673D93618CC36C CRC64;

Query Match 66.1%; Score 37; DB 12; Length 305;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
Db 225 DSGDTFFYGLV 234

Search completed: June 19, 2002, 08:29:35
Job time: 382 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:29:55 ; Search time 13.51 Seconds  
(without alignments)  
28.660 Million cell updates/sec

Title: US-09-713-299B-2  
Perfect score: 56  
Sequence: 1 DTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	1	TKS2_AEDAE
2	51	91.1	10	1	TKS1_AEDAE
3	41	73.2	2206	1	POLG_CXA21
4	39	69.6	10	1	TKN1_SCYCA
5	39	69.6	2214	1	POLG_CXA24
6	37	66.1	14	1	TKNM_RANMA
7	37	66.1	330	1	1A1D_PVRAB
8	36	64.3	11	1	TKNA_RANCA
9	36	64.3	381	1	ARGE_BUCAI
10	36	64.3	416	1	SOX_ARATH
11	35	62.5	209	1	COEC_SCYCA
12	35	62.5	213	1	CAT2_ECOLI
13	35	62.5	264	1	CPCE_SYNP7
14	35	62.5	713	1	PMIP_HUMAN
15	35	62.5	759	1	ARV1_CALVI
16	34	60.7	10	1	TKNB_RANCA
17	34	60.7	11	1	TKN_PHYFU
18	34	60.7	12	1	TKN2_KASMA
19	34	60.7	244	1	UPP_TOXGO
20	34	60.7	409	1	PEXC_PICPA
21	34	60.7	482	1	E1BL_ADEL12
22	34	60.7	509	1	AR12_DROME
23	34	60.7	524	1	CP11_MOUSE
24	34	60.7	2363	1	SPCO_MOUSE
25	34	60.7	2364	1	SPCO_HUMAN
26	33	58.9	196	1	VG53_BPT4
27	33	58.9	198	1	DUSD_HUMAN
28	33	58.9	198	1	DUSD_MOUSE
29	33	58.9	210	1	UPP_DEIRA
30	33	58.9	270	1	PSFI_DROME
31	33	58.9	325	1	1A1D_PYRHO
32	33	58.9	332	1	KPPR_SYNP3
33	33	58.9	357	1	YQ02_CABEL

RESULT 1

ID	TKS2_AEDAE	STANDARD;	PRT;	10 AA.
AC	P42635;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Sialokinin II.			
OS	Aedes aegypti (Yellowfever mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;			
OC	Aedes.			
OX	NCBI_TaxID=7159;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN=ROCKEFELLER; TISSUE=Salivary gland;			
RX	MEDLINE=94105119; PubMed=8278354;			
RA	Champagne D.E.; Ribeiro J.M.C.;			
RT	"Sialokinin I and II: vasodilatory tachykinesis from the yellow fever			
RT	mosquito Aedes aegypti.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).			
CC	-1- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE			
CC	SITE OF FEEDING.			
CC	-1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.			
DR	PIR; B49581; B49581.			
DR	InterPro; IPR002040; Tachykinin.			
DR	PROSITE; PS00267; TACHYKININ; 1.			
KW	Tachykinin; Neuropeptide; Amidation.			
FT	MOD_RES 10 10			
SQ	SEQUENCE 10 AA; 1146 MW; 3DC77C6B59C33AA8 CRC64;			

Query Match 100.0%; Score 56; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DTGDKFYGLM 10
Db	1	DTGDKFYGLM 10

RESULT 2

ID	TKS1_AEDAE	STANDARD;	PRT;	10 AA.
AC	P42634;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Sialokinin I.			
OS	Aedes aegypti (Yellowfever mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;			
OC	Aedes.			
OX	NCBI_TaxID=7159;			
RN	[1]			
RP	SEQUENCE.			

```

RC STRAIN-ROCKEFELLER: TISSUE-Salivary gland;
RX MEDLINE-94105119; PubMed-8278354;
RA Champagne D.E., Ribeiro J.M.C.;
RT "Salokinin I and II: vasodilatory tachykinins from the yellow fever
RL mosquito Aedes aegypti.";
RC Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).
CC -1- FUNCTION: VASODILATORY PEPTIDE. MAY ACTIVATE MACROPHAGES AT THE
CC SITE OF FEEDING.
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
DR PIR: A49581;
DR InterPro: IPR002040; Tachykinin.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1145 MW; 3DCFDE6B59C33AA8 CRC64;

Query Match 91.1%; Score 51; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.00045;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db :|||||||
1 NTGDKFYGLM 10

RESULT 3
POLG_CXA21 STANDARD; PRT; 2206 AA.
AC P22055;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein (Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)).
OS Coxsackievirus A21 (strain Coe).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12070;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE-90063544; PubMed-2584950;
RA Hughes P.J., North C., Minor P.D., Stanway G.;
RT "The complete nucleotide sequence of coxsackievirus A21.";
RL J. Gen. Virol. 70:2943-2952(1989).
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00538; BAA00426.1;
DR PIR: A33373; GNNY21.
DR HSPSP; P21404; 1D4M.
DR MEROPS; C03.011;
DR MEROPS; C03.022;
DR InterPro: IPR000199; Cys-protease-3C.
DR InterPro: IPR003138; Pico_P1A.

```

```

DR InterPro: IPR000081; Pico_P2A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001676; Rhv.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF00073; rhv; 3.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR ProDom; PD001125; Cys-protease-3C; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Core protein; Hydrolase; Thiol protease; Myristate.
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 341 COAT PROTEIN VP2.
FT CHAIN 342 581 COAT PROTEIN VP3.
FT CHAIN 582 879 COAT PROTEIN VP1.
FT CHAIN 880 1028 CORE PROTEIN P2A.
FT CHAIN 1029 1125 CORE PROTEIN P2B.
FT CHAIN 1126 1453 CORE PROTEIN P2C.
FT CHAIN 1454 1540 CORE PROTEIN P3A.
FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1563 1745 PICORNAIN 3C.
FT CHAIN 1746 2206 MYRISTATE (BY SIMILARITY).
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 246049 MW; 537F6A13AD37723B CRC64;

```

Query Match 73.2%; Score 41; DB 1; Length 2206;

Best Local Similarity 70.0%; Pred. No. 10;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10

Db :|||||||

799 DAGDTFYGLV 808

RESULT 4

TKNL\_SCYCA

ID TKNL\_SCYCA STANDARD; PRT; 10 AA.

AC P08608;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Scyllorhinin I.

OS Scyllorhinin canicula (Spotted dogfish) (Spotted catshark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;

OC Scyllorhinidae; Scyllorhinus.

OX NCBI\_TaxID=7830;

RN [1]

RP SEQUENCE.

RX MEDLINE-86192829; PubMed-2422058;

RA Conlon J.M., Deacon C.F., O'Toole L., Thim L.;

RT "Scyllorhinin I and II: two novel tachykinins from dogfish gut.";

RL FEBS Lett. 200:111-116(1986).

RN [2]

RP SEQUENCE.

RC TISSUE=Brain;

RX MEDLINE-93292508; PubMed-7685693;

RA Waugh D., Wang Y., Hazon N., Balmant R.J., Conlon J.M.;

RT "Primary structures and biological activities of substance-P-related

peptides from the brain of the dogfish, Scyllorhinus canicula.";

RL Eur. J. Biochem. 214:469-474(1993).

CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,

EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND

SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH



```

CC CC MUSCLES.
CC CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR DR PIR; A24867; A24867.
DR DR PIR; S33301; S33301.
DR DR InterPro; IPR002040; Tachykinin.
DR DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1219 MW; D0602D6B59C33AA9 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10
DB 4 DKFYGLM 10
| | | | | | | |
| | | | | | | |

RESULT 5
POLG_CXA24 STANDARD; PRT; 2214 AA.
AC P36290;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP1 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP4 (P1A); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A24 (strain EH24/70).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC NCBI_TaxID=36404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271460; PubMed=1317075;
RA Supanaranond K., Takeda N., Yamazaki S.;
RT "The complete nucleotide sequence of a variant of Coxsackievirus A24, an agent causing acute hemorrhagic conjunctivitis.";
RL Virus Genes 6:149-158(1992).
CC CC -!- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR DR EMBL; D90457; -; NOT_ANNOTATED_CDS.
DR DR PIR; A48548; A48548.
DR DR HSP; P03299; 1A87.
DR DR MEROPS; C03.011; -.
DR DR MEROPS; C03.022; -.
DR DR InterPro; IPR000199; Cys-protease-3C.
DR DR InterPro; IPR003138; Pico_P1A.
DR DR InterPro; IPR000081; Pico_P2A.
DR DR InterPro; IPR002527; Pico_P2B.
DR DR InterPro; IPR000605; RNA_helicase.
DR DR InterPro; IPR001205; RNA_pol_P3D.
DR DR InterPro; IPR001676; Rhv.

```

```

DR DR Pfam; PF00548; Cys-protease-3C; 1.
DR DR Pfam; PF02226; Pico_P1A; 1.
DR DR Pfam; PF00947; Pico_P2A; 1.
DR DR Pfam; PF01552; Pico_P2B; 1.
DR DR Pfam; PF00073; Rhv; 3.
DR DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR DR Pfam; PF00910; RNA_helicase; 1.
DR DR ProDom; PD001125; Cys-protease-3C; 1.
DR DR ProDom; PD001274; Pico_P2B; 1.
DR DR ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 1 69
FT CHAIN 70 340
FT CHAIN 341 577
FT CHAIN 578 888
FT CHAIN 889 1035
FT CHAIN 1036 1132
FT CHAIN 1133 1461
FT CHAIN 1462 1548
FT CHAIN 1549 1570
FT CHAIN 1571 1753
FT CHAIN 1754 2214
FT LIPID 2
FT ACT_SITE 1717 1717
FT ACT_SITE 1731 1731
SQ SEQUENCE 2214 AA; 247212 MW; E0DD74569E1B22B8 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 2214;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
DB 805 DSGDTTYGLV 814
| : | | : | | | :
| : | | : | | | :

RESULT 6
TKNM_RANMA STANDARD; PRT; 14 AA.
AC P40951;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ranamargarin.
OS Rana margaratae (Chinese frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=121156;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=90026852; PubMed=2803524;
RA Tang Y.Q., Tian S.H., Wu S.X., Hua J.C., Wu G.F., Zhao E.M., Lu Y.A.,
RA Zhu Y.Q., Zhou G., Tsou K.;
RT "Isolation and structure of ranamargarin, a new tachykinin from the skin of Chinese frog Rana margaratae.";
RL Sci. China, B, Chem. Life Sci. 32:570-579(1989).
RN [2]
RP SYNTHESIS.
RX MEDLINE=90253600; PubMed=2340087;
RA Lu Y.A., Peng J.L., Zhu Y.Q., Wu S.X., Tang Y.Q., Tian S.H., Zou G.;
RT "Synthesis and biological activity of a new frog skin peptide, ranamargarin.";
RL Sci. China, B, Chem. Life Sci. 33:170-177(1990).
CC CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS, EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH MUSCLES.
CC CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR DR InterPro; IPR002040; Tachykinin.
DR DR PROSITE; PS00267; TACHYKININ; 1.

```

KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
 FT MOD\_RES 14 14 AMIDATION  
 SQ SEQUENCE 14 AA; 1617 MW; D4593AE408C3673D CRC64;

Query Match 66.1%; Score 37; DB 1; Length 14;  
 Best Local Similarity 70.0%; Pred. No. 0.32;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10  
 I | | | | |  
 Db 5 DRAKKFYGLM 14

RESULT 7  
 LAID\_PYRAB STANDARD; PRT; 330 AA.  
 AC Q9V2L2;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative 1-aminocyclopropane-1-carboxylate deaminase (EC 4.1.99.4)  
 DE (ACC deaminase).  
 GN PAB2303.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ORSAY;  
 RA Heilig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: 1-aminocyclopropane-1-carboxylate + H(2)O = 2-oxobutanate + NH(3).  
 CC  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE ACC DEAMINASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AJ248283; CAB48986.1; -.  
 DR InterPro: IPR001926; PALP.  
 DR Pfam: PF00291; PALP; 1.  
 KW Hypothetical protein; Lyase; Pyridoxal phosphate; Complete proteome.  
 FT BINDING 54 54 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 SQ SEQUENCE 330 AA; 35754 MW; CC8699923C0B11CD CRC64;

Query Match 66.1%; Score 37; DB 1; Length 330;  
 Best Local Similarity 77.8%; Pred. No. 8.5;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10  
 I | | | | |  
 Db 287 TGRAFYGLM 295

RESULT 8  
 TKNA\_RANCA STANDARD; PRT; 11 AA.  
 AC P22688;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ranatachykinin A (RTK A).  
 OS Rana catesbeiana (Bull frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Brain, and Intestine;  
 RX MEDLINE=91254337; PubMed=2043143;  
 RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;  
 RT "Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and intestine.";  
 RL Biochem. Biophys. Res. Commun. 177:588-595(1991).  
 RN [2]  
 RP SEQUENCE  
 RC TISSUE=Intestine;  
 RX MEDLINE=94023216; PubMed=8210506;  
 RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;  
 RT "Four novel tachykinins in frog (Rana catesbeiana) brain and intestine.";  
 RL Regul. Pept. 46:81-88(1993).

CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS, EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH MUSCLES.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR: JE0426; JE0426.  
 DR PIR: A61033; A61033.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Amidation.  
 KW MOD\_RES 11 11 AMIDATION.  
 FT MOD\_RES 11 11  
 SQ SEQUENCE 11 AA; 1311 MW; 200D60CC59D40AB7 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 11;  
 Best Local Similarity 85.7%; Pred. No. 0.39;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
 I | | | | |  
 Db 5 DRFYGLM 11

RESULT 9  
 ARGE\_BUCAI STANDARD; PRT; 381 AA.  
 ID ARGE\_BUCAI  
 AC P57155;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Acetylornithine deacetylase (EC 3.5.1.16) (Acetylornithinase) (AO) (N-acetylornithinase) (NAO).  
 DE ARGE OR BU047.  
 GN Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC -1- CATALYTIC ACTIVITY: N2-acetyl-L-ornithine + H(2)O = acetate + L-ornithine.  
 CC  
 CC -1- COFACTOR: COBALT AND GLUTATHIONE (BY SIMILARITY).  
 CC  
 CC -1- PATHWAY: FIFTH STEP IN ARGININE BIOSYNTHESIS.  
 CC  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M20B; ALSO KNOWN AS THE

```
CC ARGE/DAPE/ACYL/CPG2/VSCS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BAB12770.1; -
DR InterPro; IPR001261; ARGE_DAPE_CPG2.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
DR PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.
DR PROSITE; PS00759; ARGE_DAPE_CPG2_2; FALSE_NEG.
KW Arginine biosynthesis; Hydrolase; Cobalt; Complete proteome.
SQ SEQUENCE 381 AA; 43045 MW; A837B49131A1E0ED CRC64;

Query Match 64.3%; Score 36; DB 1; Length 381;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
DB 99 ETNNKFYGL 107

RESULT 10
SOX_ARATH STANDARD; PRT; 416 AA.
ID SOX_ARATH
AC Q9SJA7;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Potential sarcosine oxidase (EC 1.5.3.1).
GN AT224580 OR F25P17.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RA thaliana.";
RL Nature 402:761-768(1999).
CC -!- CATALYTIC ACTIVITY: Sarcosine + H(2)O + O(2) -> glycine +
CC formaldehyde + H(2)O(2).
CC -!- COFACTOR: FAD; 1 mole of FAD per mole of enzyme (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MSOX/MFOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC006954; AAD23888.1; -
DR HSP; P40859; 1B3M.
```

```
DR InterPro; IPR000205; NAD_binding.
KW Hypothetical protein; Oxidoreductase; FAD; Flavoprotein.
FT NP_BIND 10 40 FAD (ADP PART) (POTENTIAL).
FT NP_BIND 325 325 FAD (COVALENT) (PROBABLE).
SQ SEQUENCE 416 AA; 45702 MW; E12E3792DF305BCC CRC64;

Query Match 64.3%; Score 36; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDKFYG 8
DB 197 GDKFYG 202

RESULT 11
COEC_SCYCA STANDARD; PRT; 209 AA.
ID COEC_SCYCA
AC P81130;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Egg case collagen (Fragments).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=99003514; PubMed=9784402;
RA Luong T.-T., Boutillon M.-M., Garrone R., Knight D.P.;
RT "Characterization of selachian egg case collagen.";
RL Biochem. Biophys. Res. Commun. 250:657-663(1998).
CC -!- FUNCTION: MAJOR COMPONENT OF THE EGG CASE WALL WHICH IS SECRETED
CC BY THE OVIDUCT. THE EGG CASE COMBINES MECHANICAL STRENGTH AND
CC TOUGHNESS WITH HIGH PERMEABILITY TO SMALL MOLECULES AND IONS.
CC -!- CAUTION: THE ORDER OF THE PEPTIDES IS UNKNOWN.
DR InterPro; IPR000087; Collagen.
KW Connective tissue; Repeat; Collagen.
FT DOMAIN 1 >129 NONHELICAL REGION.
FT DOMAIN <130 >209 TRIPLE-HELICAL REGION.
FT NON_CONS 40 41
FT NON_CONS 67 68
FT NON_CONS 94 95
FT NON_CONS 129 130
FT NON_CONS 144 145
FT NON_CONS 157 158
FT NON_CONS 177 178
FT NON_TER 209 209
SQ SEQUENCE 209 AA; 20997 MW; 7ABFFA289C5D6A45 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 209;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
DB 54 DDGKSFYGLL 63

RESULT 12
CAT2_ECOLI STANDARD; PRT; 213 AA.
ID CAT2_ECOLI
AC P22615;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Chloramphenicol acetyltransferase II (EC 2.3.1.28) (CAT-II).
```

GN CMLA.  
 OS Escherichia coli.  
 OG Plasmid Incw p8a.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-K12;  
 RX MEDLINE=91097467; PubMed=2268278;  
 RA Murray I.A., Martinez-Suarez J.V., Close T.J., Shaw W.V.;  
 RT "Nucleotide sequences of genes encoding the type II chloramphenicol  
 acetyltransferases of Escherichia coli and Haemophilus influenzae,  
 which are sensitive to inhibition by thiol-reactive reagents.";  
 RL Biochem. J. 272:505-510(1990).  
 CC -!- FUNCTION: THIS ENZYME IS AN EFFECTOR OF CHLORAMPHENICOL RESISTANCE  
 IN BACTERIA.  
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + chloramphenicol = CoA +  
 chloramphenicol 3-acetate.  
 CC -!- SUBUNIT: HOMOTRIMER.  
 CC -!- MISCELLANEOUS: TYPE II CHLORAMPHENICOL ACETYLTRANSFERASES ARE  
 SENSITIVE TO INHIBITION BY THIOL-REACTIVE REAGENTS. THE  
 INACTIVATION OCCURS AS A RESULT OF CHEMICAL MODIFICATION OF  
 CYS-26.  
 CC -!- SIMILARITY: BELONGS TO THE CHLORAMPHENICOL ACETYLTRANSFERASE  
 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X53796; CAA37805.1; -.  
 DR PIR; S13398; S13398.  
 DR HSP; P00484; IQCA.  
 DR InterPro; IPR001707; CAT.  
 DR Pfam; PF00302; CAT; 1.  
 DR ProDom; PD002660; CAT; 1.  
 DR PROSITE; PS00100; CAT; 1.  
 DR Antibiotic resistance; Transferase; Acyltransferase; Plasmid.  
 FT ACT\_SITE 189 189 BY SIMILARITY.  
 SQ SEQUENCE 213 AA; 24778 MW; 2E0D7F150A730AF7 CRC64;  
 -----  
 Query Match 62.5%; Score 35; DB 1; Length 213;  
 Best Local Similarity 70.08; Pred. No. 13;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 1 DTGDKFYGLM 10  
 Db 45 ETGYKFPPLM 54  
 -----  
 RESULT 13  
 CPCE\_SYNPF  
 ID CPCE\_SYNPF STANDARD; PRT; 264 AA.  
 AC Q44115;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Phycocyanobilin lyase alpha subunit (EC 4.1.1.1) (Phycocyanin operon  
 protein cpce).  
 GN CPCE.  
 OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=1140;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95036003; PubMed=7524727;  
 RA Bhalarao R.P., Lind L.K., Gustafsson P.;

RT "Cloning of the cpce and cpceF genes from Synechococcus sp. PCC 6301  
 and their inactivation in Synechococcus sp. PCC 7942.";  
 RL Plant Mol. Biol. 26:313-326(1994).  
 CC -!- FUNCTION: REQUIRED FOR THE CHROMOPHORYLATION OF THE CPCE GENE  
 PRODUCT.  
 CC -!- SUBUNIT: CPCE AND CPCEF ASSOCIATES TO FORM A LYASE.  
 CC -!- SIMILARITY: BELONGS TO THE CPCE/RPCE/PECE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M94218; AAA64533.1; -.  
 DR InterPro; IPR004155; HEAT\_PBS.  
 DR Pfam; PF03130; HEAT\_PBS; 5.  
 KW Phycobillinsome; Lyase.  
 SQ SEQUENCE 264 AA; 28828 MW; 49CCBA7B6321BDBF CRC64;  
 -----  
 Query Match 62.5%; Score 35; DB 1; Length 264;  
 Best Local Similarity 71.4%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 QY 1 DTGDKFY 7  
 Db 25 DTGDRYV 31  
 -----  
 RESULT 14  
 PMIP\_HUMAN  
 ID PMIP\_HUMAN STANDARD; PRT; 713 AA.  
 AC Q99797;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Mitochondrial intermediate peptidase, mitochondrial precursor  
 (EC 3.4.24.59) (MIP).  
 GN MIPEP OR MIP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=97230465; PubMed=9073519;  
 RA Chew A., Buck E.A., Peretz S., Sirugo G., Rinaldo P., Isaya G.;  
 RT "Cloning, expression, and chromosomal assignment of the human  
 mitochondrial intermediate peptidase gene (MIPEP).";  
 RL Genomics 40:493-496(1997).  
 CC -!- FUNCTION: CLEAVES PROTEINS, IMPORTED INTO THE MITOCHONDRION, TO  
 THEIR MATURE SIZE.  
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal octapeptide as second  
 stage of processing of some proteins imported in the  
 mitochondrion.  
 CC -!- ENZYME REGULATION: ACTIVITY IS DIVALENT CATION-DEPENDENT. IT IS  
 STIMULATED BY MANGANESE, MAGNESIUM OR CALCIUM IONS AND REVERSIBLY  
 INHIBITED BY ZINC, COBALT, AND IRON (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```
CC -----
DR EMBL: U80034; AAC51231.1; -.
DR MEROPS; M03.006; -.
DR MIM; 602241; -.
DR InterPro; IPR001567; Peptidase_M3.
DR InterPro; IPR00130; Zn_MTpeptidse.
DR Pfam; PF01432; Peptidase_M3.1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Transit peptide; Mitochondrion;
KW Magnesium; Manganese; Calcium; Cobalt; Iron.
FT TRANSIT 1 35 MITOCHONDRION.
FT CHAIN 36 713 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 495 495 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 496 496 BY SIMILARITY.
FT METAL 499 499 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 502 502 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 713 AA; 80611 MW; A491FCA95CD84054 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 713;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
   : : : |||||
Db 602 ETQERFYGL 610

RESULT 15
ARY1_CALVI
ID ARY1_CALVI STANDARD; PRT; 759 AA.
AC P28513; Q23815;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arylphorin subunit A4 precursor.
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=7373;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fat body;
RX MEDLINE=91282793; PubMed=1711849;
RA Naumann U., Scheller K.;
RT "Complete cDNA and gene sequence of the developmentally regulated
RT arylphorin of calliphora vicina and its homology to insect hemolymph
RT proteins and arthropod hemocyanins."
RL Biochem. Biophys. Res. Commun. 177:963-972(1991).
CC -!- FUNCTION: ARYLPHORIN IS A LARVAL STORAGE PROTEIN (LSP) WHICH MAY
CC SERVE AS A STORAGE PROTEIN USED PRIMARILY AS A SOURCE OF AROMATIC
CC AMINO ACIDS FOR PROTEIN SYNTHESIS DURING METAMORPHOSIS. IT IS A
CC CONSTITUENT OF THE SCLEROTIZING SYSTEM OF THE CUTICLE, AND SERVES
CC AS A CARRIER FOR ECYSTEROID HORMONE.
CC -!- SUBUNIT: HETEROHEXAMER.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: FAT BODY.
CC -!- SIMILARITY: TO OTHER ARYLPHORINS AND TO ARTHROPOD HEMOCYANINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76480; -; NOT_ANNOTATED_CDS.
DR EMBL: X59391; CAA42034.1; -.
DR PIR: JQ1045; JQ1045.
DR HSP: P04253; ILIA.
DR InterPro; IPR000896; Hemocyanin.
```

```
DR Pfam; PF00372; hemocyanin; 1.
DR PRINTS; PR00187; HEMOCYANIN.
DR PROSITE; PS00209; HEMOCYANIN_1; FALSE_NEG.
DR PROSITE; PS00210; HEMOCYANIN_2; 1.
KW Signal; Storage protein; Glycoprotein; Multigene family.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 759 ARYLPHORIN SUBUNIT A4.
FT CONFLICT 231 231 N -> D (IN REF. 1; CAA42034).
FT CONFLICT 351 351 N -> H (IN REF. 1; CAA42034).
FT CONFLICT 607 607 D -> E (IN REF. 1; CAA42034).
FT CONFLICT 683 683 G -> E (IN REF. 1; CAA42034).
SQ SEQUENCE 759 AA; 92342 MW; 3D222108A38B4BBD CRC64;

Query Match 62.5%; Score 35; DB 1; Length 759;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYG 8
   : : : |||||
Db 737 DTFDKYG 744

Search completed: June 19, 2002, 08:29:55
Job time: 342 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:29:35 ; Search time 41.44 Seconds  
(without alignments)  
41.746 Million cell updates/sec

Title: US-09-713-299B-2

Perfect score: 56

Sequence: 1 DTGDKFYGLM 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	11	5 Q9UAR8	Q9uar8 aedes aegyp
2	51	91.1	85	5 Q9TW63	Q9tw63 aedes aegyp
3	51	91.1	85	5 Q9UAR7	Q9uar7 aedes aegyp
4	42	75.0	304	12 Q9YLN6	Q9yln6 human coxa
5	42	75.0	305	12 Q9YLM6	Q9ylm6 human coxa
6	42	75.0	305	12 Q9YLN6	Q9yln6 human coxa
7	41	73.2	138	12 Q9YLN6	Q9yln6 human coxa
8	41	73.2	139	12 Q9YLN6	Q9yln6 human coxa
9	41	73.2	140	12 Q9YLN6	Q9yln6 human coxa
10	41	73.2	140	12 Q9YLN6	Q9yln6 human coxa
11	41	73.2	140	12 Q9YLN6	Q9yln6 human coxa
12	39	69.6	139	12 Q9YLN6	Q9yln6 human coxa
13	38	67.9	304	12 Q9YLN6	Q9yln6 human coxa
14	37	66.1	257	10 Q9MIF6	Q9mif6 arabidopsis
15	37	66.1	435	10 Q942T2	Q942t2 oryza sativ
16	37	66.1	539	16 Q9HXT2	Q9hxt2 pseudomonas

17	37	66.1	979	3	P87152	P87152 schizosacch
18	37	66.1	1134	17	O30078	O30078 archaeoglob
19	37	66.1	1157	16	Q97F62	Q97f62 clostridium
20	37	66.1	1217	16	Q97GW3	Q97gw3 clostridium
21	36	64.3	315	2	Q9RJ00	Q9rj00 streptomyce
22	36	64.3	333	2	O33560	O33560 rhodobacter
23	36	64.3	387	17	Q9HIE7	Q9hie7 thermoplasma
24	36	64.3	416	10	Q9SJA7	Q9sja7 arabidopsis
25	36	64.3	508	2	Q9R6V5	Q9r6v5 synechococc
26	36	64.3	649	16	O83749	O83749 treponema p
27	36	64.3	697	13	Q98TV9	Q98tv9 notothenia
28	36	64.3	697	13	Q98TV4	Q98tv4 chaenocarpa
29	36	64.3	697	13	Q98TV3	Q98tv3 goblionototh
30	36	64.3	766	3	Q96V54	Q96v54 emericella
31	35	62.5	109	16	Q55186	Q55186 synechocyst
32	35	62.5	171	10	Q9M3H7	Q9m3h7 cicler arlet
33	35	62.5	244	12	Q994Y5	Q994y5 sacbrood vi
34	35	62.5	244	12	Q994Y4	Q994y4 sacbrood vi
35	35	62.5	244	12	Q994Y3	Q994y3 sacbrood vi
36	35	62.5	244	12	Q994Y2	Q994y2 sacbrood vi
37	35	62.5	244	12	Q994Y1	Q994y1 sacbrood vi
38	35	62.5	244	12	Q994Y0	Q994y0 sacbrood vi
39	35	62.5	244	12	Q994X9	Q994x9 sacbrood vi
40	35	62.5	244	12	Q994X8	Q994x8 sacbrood vi
41	35	62.5	244	12	Q994X7	Q994x7 sacbrood vi
42	35	62.5	244	12	Q994X6	Q994x6 sacbrood vi
43	35	62.5	244	12	Q994X5	Q994x5 sacbrood vi
44	35	62.5	244	12	Q994X4	Q994x4 sacbrood vi
45	35	62.5	244	12	Q994X3	Q994x3 sacbrood vi

#### ALIGNMENTS

RESULT 1

Q9UAR8 ID Q9UAR8 PRELIMINARY: PRT; 11 AA.  
AC Q9UAR8:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE SIALOKININ I PREPROTEIN (FRAGMENT).  
OS Aedes aegypti (Yellow fever mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;  
OC Aedes.  
CX NCBI\_TaxID=7159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ROCKEFELLER/RED; TISSUE=SALIVARY GLAND;  
RX MEDLINE=20099025; PubMed=10620041;  
RA Beerntsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;  
RT "Characterization of the Sialokinin I gene encoding the salivary  
RT vasodilator of the yellow fever mosquito, Aedes aegypti.";  
RL Insect Mol. Biol. 8:459-467(1999).  
DR EMBL; AF108100; AAD16884.1; -  
DR InterPro; IPR02040; Tachykinin.  
DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_1.  
FT NON\_TER 1  
SQ SEQUENCE 11 AA; 1203 MW; 8BADC77C6B59C33A CRC64;

Query Match 100.0%; Score 56; DB 5; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTGDKFYGLM 10

Db 1 DTGDKFYGLM 10

RESULT 2

Q9TW63

```
ID Q9TWG3 PRELIMINARY; PRT; 85 AA.
AC Q9TWG3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE STALOKININ I PREPROTEIN.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ROCKEFELLER;
RA Beertsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT "Characterization of the sialokinin I gene encoding the salivary
vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL Insect Mol. Biol. 0:0-0(1999).
DR EMBL; AF108102; AAD16886.1; -.
DR EMBL; AF108101; AAD16885.1; -.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
SQ SEQUENCE 85 AA; 9326 MW; 209F03E3932A6A8B CRC64;

Query Match 91.13; Score 51; DB 5; Length 85;
Best Local Similarity 90.08; Pred. No. 0.023;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 75 NTGDKFYGLM 84
:|||||

RESULT 3
Q9UAR7 PRELIMINARY; PRT; 85 AA.
AC Q9UAR7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SALIVARY VASODILATORY PROTEIN PRECURSOR.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ROCKEFELLER;
RA Beertsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT "Characterization of the sialokinin I gene encoding the salivary
vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL Insect Mol. Biol. 8:459-467(1999).
DR EMBL; AF108099; AAD17916.1; -.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
KW SIGNAL.
FT CHAIN 1 23 POTENTIAL..
FT SIGNAL 75 85 SALIVARY VASODILATORY PROTEIN.
FT CHAIN 85 AA; 9460 MW; ED4669E3470638DF CRC64;

Query Match 91.13; Score 51; DB 5; Length 85;
Best Local Similarity 90.08; Pred. No. 0.023;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 75 NTGDKFYGLM 84
:|||||

RESULT 4
Q9YLM6 PRELIMINARY; PRT; 304 AA.
AC Q9YLM6;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CAPSID PROTEIN VPI (FRAGMENT).
GN VPI.
OS Human coxsackievirus A11.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=42770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BELGIUM-1;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
serotype with VPI sequence and application to picornavirus
classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081301; AAD17705.1; -.
DR HSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1 1
FT NON_TER 304 304
SQ SEQUENCE 304 AA; 33884 MW; F151A7C24CD7A4C1 CRC64;

Query Match 75.0%; Score 42; DB 12; Length 304;
Best Local Similarity 77.8%; Pred. No. 4.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
Db 225 DTGDAYGL 233
|||||

RESULT 5
Q9YLM6 PRELIMINARY; PRT; 305 AA.
AC Q9YLM6;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CAPSID PROTEIN VPI (FRAGMENT).
GN VPI.
OS Human coxsackievirus A24.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12089;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JOSEPH;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
serotype with VPI sequence and application to picornavirus
classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081311; AAD17715.1; -.
DR HSP; P03300; 2PLV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1 1
FT NON_TER 305 305
SQ SEQUENCE 305 AA; 34091 MW; D4673D93618CC36C CRC64;

Query Match 75.0%; Score 42; DB 12; Length 305;
Best Local Similarity 70.0%; Pred. No. 4.9;
```

```

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 DTGDKFYGLM 10
    I:| | | | |
Db 225 DSGDTFYGLV 234

RESULT 6
Q9YLLJ4
ID Q9YLLJ4 PRELIMINARY; PRT; 305 AA.
AC Q9YLLJ4;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Echovirus 34.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=86108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DN-19;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081347; AAD17751.1; -.
DR HSSP; Q84790; LPVC.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1 305
FT NON_TER 305 305
SQ SEQUENCE 305 AA; 34223 MW; CE00DEC18490B6E2 CRC64;

Query Match 75.0%; Score 42; DB 12; Length 305;
Best Local Similarity 70.0%; Pred. No. 4.9;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 DTGDKFYGLM 10
    I:| | | | |
Db 225 DSGDTFYGLV 234

RESULT 7
Q9JA93
ID Q9JA93 PRELIMINARY; PRT; 138 AA.
AC Q9JA93;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A21.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD86-7277;
RX MEDLINE=20164691; PubMed=10699015;
RA Oberste M.S., Maher K., Flemister M.R., Marchetti G., Kilpatrick D.R.,
RA Pallansch M.A.;
RT "Comparison of Classic and Molecular Approaches for the Identification
RT of Untypeable Enteroviruses.";
RL J. Clin. Microbiol. 38:1170-1174(2000).
DR EMBL; AF152265; AAF43560.1; -.
DR HSSP; P03300; LPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.

```

```

FT NON_TER 1 138
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15473 MW; 310A95CDAE7200A8 CRC64;

Query Match 73.2%; Score 41; DB 12; Length 138;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 DTGDKFYGLM 10
    I:| | | | |
Db 59 DAGDTFYGLV 68

RESULT 8
Q9W902
ID Q9W902 PRELIMINARY; PRT; 139 AA.
AC Q9W902;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VP1 (FRAGMENT).
OS Human coxsackievirus A21.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TX96-2184; AZ94-2060, AND GA95-2093;
RX MEDLINE=99221739; PubMed=10203472;
RA Oberste M.S., Maher K., Kilpatrick D.R., Flemister M.R., Brown B.A.,
RA Pallansch M.A.;
RT "Typing of human enteroviruses by partial sequencing of VP1.";
RL J. Clin. Microbiol. 37:1288-1293(1999).
DR EMBL; AF081637; AAD40541.1; -.
DR EMBL; AF081598; AAD40502.1; -.
DR EMBL; AF081612; AAD40516.1; -.
DR HSSP; P03300; LPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1 139
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15621 MW; 8A110A95CDAE7200 CRC64;

Query Match 73.2%; Score 41; DB 12; Length 139;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 DTGDKFYGLM 10
    I:| | | | |
Db 59 DAGDTFYGLV 68

RESULT 9
Q9JAA0
ID Q9JAA0 PRELIMINARY; PRT; 140 AA.
AC Q9JAA0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A21.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUT88-8020;
RX MEDLINE=20164691; PubMed=10699015;
RA Oberste M.S., Maher K., Flemister M.R., Marchetti G., Kilpatrick D.R.,
RA Pallansch M.A.;

```



```
RT "Comparison of Classic and Molecular Approaches for the Identification
RT of Untypeable Enteroviruses.";
RL J. Clin. Microbiol. 38:1170-1174(2000);
DR EMBL; AF152258; AAF43553.1; -.
DR HSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 140
SQ SEQUENCE 140 AA; 15664 MW; B8F7ADC79554AE77 CRC64;

Query Match 73.2%; Score 41; DB 12; Length 140;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 59 DAGDTFYGLV 68

RESULT 10
Q9JA99 PRELIMINARY; PRT; 140 AA.
AC Q9JA99;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A21.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GUT88-8438;
RX MEDLINE-20164691; PubMed=10699015;
RA Oberste M.S., Maher K., Flemister M.R., Marchetti G., Kilpatrick D.R.,
RA Pallansch M.A.;
RT "Comparison of Classic and Molecular Approaches for the Identification
RT of Untypeable Enteroviruses.";
RL J. Clin. Microbiol. 38:1170-1174(2000).
DR EMBL; AF152259; AAF43554.1; -.
DR HSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 140
SQ SEQUENCE 140 AA; 15664 MW; B8F7ADC79554AE77 CRC64;

Query Match 73.2%; Score 41; DB 12; Length 140;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 59 DAGDTFYGLV 68

RESULT 11
Q9JA58 PRELIMINARY; PRT; 140 AA.
AC Q9JA58;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A21.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
```

```
OX NCBI_TaxID=12069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WA89-9165;
RX MEDLINE-20164691; PubMed=10699015;
RA Oberste M.S., Maher K., Flemister M.R., Marchetti G., Kilpatrick D.R.,
RA Pallansch M.A.;
RT "Comparison of Classic and Molecular Approaches for the Identification
RT of Untypeable Enteroviruses.";
RL J. Clin. Microbiol. 38:1170-1174(2000).
DR EMBL; AF152300; AAF43595.1; -.
DR HSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 140
SQ SEQUENCE 140 AA; 15678 MW; 689A110A95CDAE72 CRC64;

Query Match 73.2%; Score 41; DB 12; Length 140;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 59 DAGDTFYGLV 68

RESULT 12
Q9WRY4 PRELIMINARY; PRT; 139 AA.
AC Q9WRY4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP1 (FRAGMENT).
OS Human coxsackievirus A24.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12089;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DOR93-1657;
RX MEDLINE-99221739; PubMed=10203472;
RA Oberste M.S., Maher K., Kilpatrick D.R., Flemister M.R., Brown B.A.,
RA Pallansch M.A.;
RT "Typing of human enteroviruses by partial sequencing of VP1.";
RL J. Clin. Microbiol. 37:1288-1293(1999).
DR EMBL; AF081603; AAD40507.1; -.
DR HSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 15780 MW; F66EA00A51D23D91 CRC64;

Query Match 69.6%; Score 39; DB 12; Length 139;
Best Local Similarity 60.0%; Pred. No. 8;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10
Db 59 DSGDTFYGLV 68

RESULT 13
Q9YLN2 PRELIMINARY; PRT; 304 AA.
AC Q9YLN2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
```

```
DE CAPSID PROTEIN VP1 (FRAGMENT).
GN VP1.
OS Human coxsackievirus A15.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=42774;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G-9;
RX MEDLINE=99138973; PubMed=9971773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification.";
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081305; RAD17709.1; -.
DR HSSP; P03300; IPOV.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 304
SQ SEQUENCE 304 AA; 33899 MW; 03B5D55F160F49B5 CRC64;

Query Match 67.9%; Score 38; DB 12; Length 304;
Best Local Similarity 66.7%; Pred. NO. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
Db 225 DSGDAYYGL 233

RESULT 14
Q9M1F6 PRELIMINARY; PRT; 257 AA.
AC Q9M1F6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 29.7 KDA PROTEIN.
GN F9K21.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordan N., Bangert S., Wiedemann R., Voss H., Unseld M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL138657; CAB75480.1; -.
DR InterPro; IPR002867; IBR.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01485; IBR; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW Hypothetical protein; Zinc-finger.
SQ SEQUENCE 257 AA; 29728 MW; 57EF280C649E4A95 CRC64;

Query Match 66.1%; Score 37; DB 10; Length 257;
Best Local Similarity 66.7%; Pred. NO. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
```

```
Db 240 ETGGRFYGL 248

RESULT 15
Q942T2 PRELIMINARY; PRT; 435 AA.
AC Q942T2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P0583G08.7 PROTEIN.
GN P0583G08.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0583G08.7";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003282; BAB64772.1; -.
SQ SEQUENCE 435 AA; 47559 MW; 49CB0F59D6B46DF5 CRC64;

Query Match 66.1%; Score 37; DB 10; Length 435;
Best Local Similarity 85.7%; Pred. NO. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGDKFYGL 8
Db 195 TGDQFYGL 201

Search completed: June 19, 2002, 08:29:36
Job time: 383 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:20:23 ; Search time 26.86 Seconds  
(without alignments)  
35.774 Million cell updates/sec

Title: US-09-713-299B-1  
Perfect score: 56  
Sequence: 1 NTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Prqd. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	2 A49581	sialokinin I - yel
2	51	91.1	10	2 B49581	sialokinin II - ye
3	42	75.0	1157	2 F97255	fusion of alpha-gl
4	42	75.0	1217	2 F97177	alpha-glucosidase
5	39	69.6	10	2 A24867	scyllorhinin I - s
6	38	67.9	415	2 B97085	DNA-methyltransfer
7	38	67.9	845	1 GLVK	beta-glucosidase (
8	38	67.9	867	1 C64785	outer membrane ush
9	38	67.9	869	2 B90703	probable outer mem
10	38	67.9	869	2 E85553	hypothetical prote
11	37	66.1	330	2 C75192	probable l-aminoc
12	37	66.1	478	2 C81103	arginate O-acetyla
13	37	66.1	478	2 H81838	probable polysacch
14	37	66.1	539	2 E83181	probable MFS trans
15	37	66.1	572	1 VGNZBA	cell fusion glycop
16	37	66.1	574	1 VGNZBS	cell fusion glycop
17	37	66.1	574	2 JQ1583	cell fusion protei
18	37	66.1	1134	2 G69269	molybdopterin-bind
19	36	64.3	11	2 A61033	ranacthykinin A -
20	36	64.3	205	2 C71691	ribosomal protein
21	36	64.3	205	2 D97758	30S ribosomal prot
22	36	64.3	308	2 AE2829	hypothetical prote
23	36	64.3	328	2 T00747	RING-H2 finger pro
24	36	64.3	333	2 T45023	hypothetical prote
25	36	64.3	337	2 C97607	hypothetical prote
26	36	64.3	393	2 S77896	guanidinooacetate k
27	36	64.3	416	2 D84638	probable sarcosine
28	36	64.3	447	2 C64828	probable polynucle
29	36	64.3	447	2 A99751	probable polynucle

30	36	64.3	447	2 G85614	probable polynucle
31	36	64.3	447	2 AH0611	conserved hypother
32	36	64.3	785	2 T38359	cullin 3 homolog -
33	36	64.3	798	2 S62405	hypothetical prote
34	36	64.3	2206	1 GNNY21	genome polyprotein
35	35	62.5	109	2 S74409	hypothetical prote
36	35	62.5	157	2 A87163	conserved hypother
37	35	62.5	161	2 AD2540	hypothetical prote
38	35	62.5	223	2 T19018	hypothetical prote
39	35	62.5	231	2 H64466	hypothetical prote
40	35	62.5	257	2 T47491	hypothetical prote
41	35	62.5	267	2 H90368	hypothetical prote
42	35	62.5	285	2 F71076	hypothetical prote
43	35	62.5	311	2 T15997	hypothetical prote
44	35	62.5	405	2 T41593	hypothetical prote
45	35	62.5	486	2 D89911	amino acid carrier

#### ALIGNMENTS

##### RESULT 1

A49581

sialokinin I - yellow fever mosquito

C:Species: Aedes aegypti (yellow fever mosquito)

C>Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 18-Aug-2000

C:Accession: A49581

R:Champagne, D.E.: Ribeiro, J.M.

Proc. Natl. Acad. Sci. U.S.A. 91, 138-142, 1994

A:Title: Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito

A:Reference number: A49581; MUID:94105119

A:Accession: A49581

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <CHA>

A:Experimental source: Rockefeller, salivary gland

A>Note: sequence extracted from NCBI backbone (NCBIP:141841)

C:Superfamily: unassigned animal peptides

Query Match 100.0%; Score 56; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTGDKFYGLM	10
Db	1	NTGDKFYGLM	10

##### RESULT 2

B49581

sialokinin II - yellow fever mosquito

C:Species: Aedes aegypti (yellow fever mosquito)

C>Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 18-Aug-2000

C:Accession: B49581

R:Champagne, D.E.: Ribeiro, J.M.

Proc. Natl. Acad. Sci. U.S.A. 91, 138-142, 1994

A:Title: Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito

A:Reference number: A49581; MUID:94105119

A:Accession: B49581

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <CHA>

A:Experimental source: Rockefeller, salivary gland

A>Note: sequence extracted from NCBI backbone (NCBIP:141842)

C:Superfamily: unassigned animal peptides

##### Query Match

91.1%; Score 51; DB 2; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0019;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTGDKFYGLM	10
----	---	------------	----

Db 1 DTGDKFYGLM 10  
:|||||

## RESULT 3

F97255 fusion of alpha-glucosidase (family 31 glycosyl hydrolase) and glycosidase (TreA/Wals fa  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: F97255  
R:; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; PMID:21359325; PMID:21359325  
A:Accession: F97255  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1157 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80833.1; PID:gl5025938; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2891

Query Match 75.0%; Score 42; DB 2; Length 1157;  
Best Local Similarity 66.7%; Pred. No. 11;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGL 9  
:|||||  
Db 165 NSGDRFYGI 173

## RESULT 4

F97177 alpha-glucosidase fused to unknown alpha-amylase C-terminal. domain [imported] - Clostr  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: F97177  
R:; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; PMID:21359325; PMID:21359325  
A:Accession: F97177  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1217 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80209.1; PID:gl5025253; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2252

Query Match 75.0%; Score 42; DB 2; Length 1217;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGL 9  
:|||||  
Db 169 NSGDRFYGI 177

## RESULT 5

A24867 scylorhinin I - smaller spotted catshark  
C:Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)  
C:Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 18-Aug-2000  
C:Accession: A24867; S33301  
R:Conlon, J.M.; Deacon, C.F.; O'Toole, L.; Thim, L.  
FEBS Lett. 200, 111-116, 1986  
A:Title: Scyliorhinin I and II: two novel tachykinins from dogfish gut.

A:Reference number: A91359; PMID:86192829  
A:Accession: A24867  
A:Molecule type: protein  
A:Residues: 1-10 <CON>  
R:; Waugh, D.; Wang, Y.; Hazon, N.; Balmert, R.J.; Conlon, J.M.  
Eur. J. Biochem. 214, 469-474, 1993  
A:Title: Primary structures and biological activities of substance-P-related peptides  
A:Reference number: S33301; PMID:93292508  
A:Accession: S33301  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <WAU>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; neuropeptide  
F:10/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 69.6%; Score 39; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
|||||  
Db 4 DKFYGLM 10

## RESULT 6

B97085 DNA-methyltransferase (cytosine-specific), ortholog of BSP6I Bsubtilis [imported] - C  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: B97085  
R:; Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; PMID:21359325; PMID:21359325  
A:Accession: B97085  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-415 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK79469.1; PID:gl5024448; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1501

Query Match 67.9%; Score 38; DB 2; Length 415;  
Best Local Similarity 85.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFY 7  
|||||  
Db 266 NTGDKYI 272

## RESULT 7

GLVK beta-glucosidase (EC 3.2.1.21) precursor - yeast (Kluyveromyces marxianus var. marxia  
N:Alternate names: beta-D-glucoside glucohydrolase; cellobiase; gentiobiase  
C:Species: Kluyveromyces marxianus var. marxianus, Candida kefyr  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 28-May-1999  
C:Accession: A29148  
R:; Raynal, A.; Gerbaud, C.; Francinques, M.C.; Guerinneau, M.  
Curr. Genet. 12, 175-184, 1987  
A:Title: Sequence and transcription of the beta-glucosidase gene of Kluyveromyces fra  
A:Reference number: A29148; PMID:88210533  
A:Accession: A29148  
A:Molecule type: DNA  
A:Residues: 1-845 <RAY>  
A:Cross-references: GB:X05918; NID:g2804; PIDN:CAA29353.1; PID:g2805  
C:Superfamily: beta-glucosidase  
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

QY	1	NTGDKFYGLM	10
Db	92	NTGDKSGGLM	101

RESULT 12  
C81103  
alginate O-  
C;Species:  
C;Date: 31.

C:Accession: C81103  
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
ri, H.; Qin, H.; Vanthaeap, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755  
A:Accession: C81103  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-478 <TET>  
A:Cross-references: GB:AE002475; GB:AE002098; NID:7226502; PIDN:AAF41650.1; PID:g722651  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1273

Query Match 66.1%; Score 37; DB 2; Length 478;  
Best Local Similarity 62.5%; Pred. No. 41;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8  
||||:|  
Db 350 NTGDRYFG 357

RESULT 13  
H81838  
probable polysaccharide modification protein NMA1478 [imported] - *Neisseria meningitidis*  
C:Species: *Neisseria meningitidis*  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: H81838  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: A81775; MUID:20222556  
A:Accession: H81838  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-478 <PAR>  
A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84711.1; PID:g738012  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA1478

Query Match 66.1%; Score 37; DB 2; Length 478;  
Best Local Similarity 62.5%; Pred. No. 41;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8  
||||:|  
Db 350 NTGDRYFG 357

RESULT 14  
E83181  
probable MFS transporter PA3709 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: E83181  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337  
A:Accession: E83181  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-539 <STO>

A:Cross-references: GB:AE004790; GB:AE004091; NID:g9949873; PIDN:AAG07096.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3709

Query Match 66.1%; Score 37; DB 2; Length 539;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKEYGL 9  
|||||  
Db 504 TGDIFVGL 511

RESULT 15  
VGNZBA

cell fusion glycoprotein precursor - bovine respiratory syncytial virus (strain A5190  
N:Contains: cell fusion glycoprotein F1; cell fusion glycoprotein F2  
C:Species: bovine respiratory syncytial virus  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999

C:Accession: JQ1481  
R:Zamora, M.; Samal, S.K.  
J. Gen. Virol. 73, 737-741, 1992  
A:Title: Sequence analysis of M2 mRNA of bovine respiratory syncytial virus obtained  
A:Reference number: JQ1481; MUID:92185490  
A:Accession: JQ1481  
A:Molecule type: mRNA  
A:Residues: 1-572 <ZAM>

A:Cross-references: GB:M82816; NID:g210823; PIDN:AAA42804.1; PID:g210824  
C:Genetics:  
A:Gene: F

C:Superfamily: cell fusion glycoprotein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-136/Product: cell fusion glycoprotein F2 #status predicted <F2P>  
F:137-572/Product: cell fusion glycoprotein F1 #status predicted <F1P>  
F:138-154/Domain: transmembrane #status predicted <TM1>  
F:524-540/Domain: transmembrane #status predicted <TM2>  
F:27-70,120,498,567/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.1%; Score 37; DB 1; Length 572;  
Best Local Similarity 70.0%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKKEYGLM 10  
|: |||||  
Db 120 NSTKKEYGLM 129

Search completed: June 19, 2002, 08:24:42  
Job time: 259 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:20:48 ; Search time 25.97 Seconds  
(without alignments)  
33.298 Million cell updates/sec

Title: US-09-713-299B-1  
Perfect score: 56  
Sequence: 1 NTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262238 seqs, 86476186 residues

Total number of hits satisfying chosen parameters: 262238

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	5	US-09-713-299B-1
2	51	91.1	10	5	US-09-713-299B-2
3	37	66.1	330	7	US-60-360-039-1233
4	37	66.1	428	5	US-09-540-209B-8953
5	36	64.3	885	7	US-60-360-039-2061
6	35	62.5	155	5	US-09-445-289B-5
7	35	62.5	182	5	US-09-445-289B-43
8	35	62.5	220	5	US-09-445-289B-36
9	35	62.5	354	7	US-60-360-039-18238
10	35	62.5	426	7	US-60-360-039-21051
11	35	62.5	3390	5	US-09-935-625-1866
12	35	62.5	3390	5	US-09-935-625-5872
13	35	62.5	3390	5	US-09-935-625-22358
14	35	62.5	3390	5	US-09-935-625-30750
15	35	62.5	3437	5	US-09-935-625-1865
16	35	62.5	3437	5	US-09-935-625-5871
17	35	62.5	3437	5	US-09-935-625-22357
18	35	62.5	3437	5	US-09-935-625-30749
19	34	60.7	488	5	US-09-540-209B-6937
20	34	60.7	1422	5	US-09-935-625-14192
21	34	60.7	1422	5	US-09-935-625-31389
22	34	60.7	1459	5	US-09-935-625-14191
23	34	60.7	1459	5	US-09-935-625-31388
24	34	60.7	1462	5	US-09-935-625-14190
25	34	60.7	1462	5	US-09-935-625-31387
26	33	58.9	81	6	US-10-004-860-740

27 33 58.9 165 5 US-09-540-209B-5572  
28 33 58.9 185 7 US-60-360-039-23006  
29 33 58.9 235 6 US-10-004-860-318  
30 33 58.9 235 6 US-10-004-860-738  
31 33 58.9 275 7 US-60-360-039-20009  
32 33 58.9 315 5 US-09-540-209B-5556  
33 33 58.9 328 7 US-60-360-039-20283  
34 33 58.9 502 7 US-60-360-039-3589  
35 33 58.9 647 7 US-60-360-039-10283  
36 33 58.9 983 7 US-60-360-039-16824  
37 32 57.1 99 5 US-09-445-289B-8  
38 32 57.1 99 5 US-09-445-289B-45  
39 32 57.1 188 5 US-09-445-289B-2  
40 32 57.1 362 5 US-09-445-289B-1  
41 32 57.1 362 7 US-60-360-039-899  
42 32 57.1 390 1 PCT-US02-13142-3573  
43 32 57.1 390 6 US-10-128-714-3573  
44 32 57.1 419 5 US-09-540-209B-7659  
45 32 57.1 433 5 US-09-540-209B-6919

ALIGNMENTS

RESULT 1  
US-09-713-299B-1  
; Sequence 1, Application US/09713299B  
; GENERAL INFORMATION:  
; APPLICANT: Maier, Nathan C.  
; APPLICANT: Jarstfer, Amiel G.  
; TITLE OF INVENTION: The Use Of Mosquito Salivary Tachkinins To Remediate Unregulat  
; TITLE OF INVENTION: Proliferation  
; FILE REFERENCE: MAI5302.01A  
; CURRENT APPLICATION NUMBER: US/09/713,299B  
; CURRENT FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Aedes aegypti  
US-09-713-299B-1

Query Match 100.0%; Score 56; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTGDKFYGLM 10  
Db 1 NTGDKFYGLM 10

RESULT 2  
US-09-713-299B-2  
; Sequence 2, Application US/09713299B  
; GENERAL INFORMATION:  
; APPLICANT: Maier, Nathan C.  
; APPLICANT: Jarstfer, Amiel G.  
; TITLE OF INVENTION: The Use Of Mosquito Salivary Tachkinins To Remediate Unregulat  
; TITLE OF INVENTION: Proliferation  
; FILE REFERENCE: MAI5302.01A  
; CURRENT APPLICATION NUMBER: US/09/713,299B  
; CURRENT FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Aedes aegypti  
US-09-713-299B-2

Query Match 91.1%; Score 51; DB 5; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0012;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
:|||||  
Db 1 DTGDKFYGLM 10

## RESULT 3

US-60-360-039-1233  
; Sequence 1233, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360,039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 1233

; LENGTH: 330

; TYPE: PRT

; ORGANISM: Pyrococcus abyssi

US-60-360-039-1233

Query Match 66.1%; Score 37; DB 7; Length 330;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10  
:|||||  
Db 287 TGAIFYGLM 295

## RESULT 4

US-09-540-209B-8953

; Sequence 8953, Application US/09540209B

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA

; FILE REFERENCE: 2709,1001-001

; CURRENT APPLICATION NUMBER: US/09/540,209B

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444

; SEQ ID NO 8953

; LENGTH: 428

; TYPE: PRT

; ORGANISM: B fragilis

US-09-540-209B-8953

Query Match 66.1%; Score 37; DB 5; Length 428;  
Best Local Similarity 77.8%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGL 9  
:|||||  
Db 105 NPGDKFMGL 113

## RESULT 5

US-60-360-039-2061

; Sequence 2061, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360,039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 2061

; LENGTH: 885

; TYPE: PRT

; ORGANISM: Schizosaccharomyces pombe

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(885)

; OTHER INFORMATION: unsure at all Xaa locations

US-60-360-039-2061

Query Match 64.3%; Score 36; DB 7; Length 885;  
Best Local Similarity 85.7%; Pred. No. 84;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFY 7  
:|||||

Db 752 NAGDKFY 758

## RESULT 6

US-09-445-289B-5

; Sequence 5, Application US/09445289B

; GENERAL INFORMATION:

; APPLICANT: Mukamolova, Galina V. et al.

; TITLE OF INVENTION: Bacterial Pheromones and Uses Therefor

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/445,289B

; FILING DATE: 11-May-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB98/01619

; FILING DATE: 03-MAY-1998

; APPLICATION NUMBER: GB 9711389.8

; FILING DATE: 04-JUN-1997

; APPLICATION NUMBER: GB 9811221.2

; FILING DATE: 27-MAY-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauro, Peter C.

; REGISTRATION NUMBER: 32,360

; REFERENCE/DOCKET NUMBER: FHW-05105

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 742-4214

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 155 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-445-289B-5



Query Match 62.5%; Score 35; DB 5; Length 155;  
Best Local Similarity 75.0%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 NTGDKFYG 8  
Db 62 NTGNGFYG 69

RESULT 7  
US-09-445-289B-43  
; Sequence 43, Application US/09445289B  
; GENERAL INFORMATION:  
; APPLICANT: Mukamolova, Galina V. et al.  
; TITLE OF INVENTION: Bacterial Pheromones and Uses Therefor  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/445,289B  
; FILING DATE: 11-May-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB98/01619  
; FILING DATE: 03-MAY-1998  
; APPLICATION NUMBER: GB 9711389.8  
; FILING DATE: 04-JUN-1997  
; APPLICATION NUMBER: GB 9811221.2  
; FILING DATE: 27-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauro, Peter C.  
; REGISTRATION NUMBER: 32,360  
; REFERENCE/DOCKET NUMBER: FHW-051US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 43:  
US-09-445-289B-43

Query Match 62.5%; Score 35; DB 5; Length 182;  
Best Local Similarity 75.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 NTGDKFYG 8  
Db 21 NTGNGFYG 28

RESULT 8  
US-09-445-289B-36  
; Sequence 36, Application US/09445289B  
; GENERAL INFORMATION:  
; APPLICANT: Mukamolova, Galina V. et al.  
; TITLE OF INVENTION: Bacterial Pheromones and Uses Therefor  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/445,289B  
FILING DATE: 11-May-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB98/01619  
FILING DATE: 03-MAY-1998  
APPLICATION NUMBER: GB 9711389.8  
FILING DATE: 04-JUN-1997  
APPLICATION NUMBER: GB 9811221.2  
FILING DATE: 27-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauro, Peter C.  
REGISTRATION NUMBER: 32,360  
REFERENCE/DOCKET NUMBER: FHW-051US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 220 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-09-445-289B-36

Query Match 62.5%; Score 35; DB 5; Length 220;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 NTGDKFYG 8  
Db 59 NTGNGFYG 66

RESULT 9  
US-60-360-039-18238  
; Sequence 18238, Application US/60360039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18238  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Thermoplasma acidophilum  
US-60-360-039-18238

Query Match 62.5%; Score 35; DB 7; Length 354;  
Best Local Similarity 77.8%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9

```
Db 202 NTMTKFYGL 210
|| |||||
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5872
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-5872

Query Match 62.5%; Score 35; DB 5; Length 3390;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9
| | | | |
Db 2 NQGNKFHGL 10
| | | | |

RESULT 13
US-09-935-625-22358
; Sequence 22358, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 22358
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-22358

Query Match 62.5%; Score 35; DB 5; Length 3390;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9
| | | | |
Db 2 NQGNKFHGL 10
| | | | |

RESULT 14
US-09-935-625-30750
; Sequence 30750, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 30750
; LENGTH: 3390
; TYPE: PRT
```

---

```
Db 202 NTMTKFYGL 210
|| |||||
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5872
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-5872

Query Match 62.5%; Score 35; DB 5; Length 3390;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9
| | | | |
Db 2 NQGNKFHGL 10
| | | | |

RESULT 12
US-09-935-625-5872
; Sequence 5872, Application US/09935625
```

---

```
Db 202 NTMTKFYGL 210
|| |||||
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5872
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-5872

Query Match 62.5%; Score 35; DB 5; Length 3390;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9
| | | | |
Db 2 NQGNKFHGL 10
| | | | |

RESULT 11
US-09-935-625-1866
; Sequence 1866, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 1866
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-1866

Query Match 62.5%; Score 35; DB 7; Length 426;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGDKPYG 8
| | | | |
Db 342 TGDDFYG 348
| | | | |

RESULT 11
US-09-935-625-1866
; Sequence 1866, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 1866
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-1866

Query Match 62.5%; Score 35; DB 7; Length 426;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGDKPYG 8
| | | | |
Db 342 TGDDFYG 348
| | | | |

RESULT 11
US-09-935-625-1866
; Sequence 1866, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 1866
; LENGTH: 3390
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..3390
; OTHER INFORMATION: Ceres Seq. ID no. 3088785
US-09-935-625-1866

Query Match 62.5%; Score 35; DB 5; Length 3390;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9
| | | | |
Db 2 NQGNKFHGL 10
| | | | |

RESULT 12
US-09-935-625-5872
; Sequence 5872, Application US/09935625
```

; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..3390  
; OTHER INFORMATION: Ceres Seq. ID no. 3088785  
US-09-935-625-30750

Query Match 62.5%; Score 35; DB 5; Length 3390;  
Best Local Similarity 66.7%; Pred. No. 5.2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9  
| | : | : |  
Db 2 NOGNKFHGL 10

RESULT 15  
US-09-935-625-1865  
; Sequence 1865, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE OF INVENTION: MODULATING VARIOUS RESPONSES  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 1865  
; LENGTH: 3437  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..3437  
; OTHER INFORMATION: Ceres Seq. ID no. 3088784  
US-09-935-625-1865

Query Match 62.5%; Score 35; DB 5; Length 3437;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYGL 9  
| | : | : |  
Db 49 NOGNKFHGL 57

Search completed: June 19, 2002, 08:25:14  
Job time: 266 sec



```
RESULT 2
US-08-900-953-2
; Sequence 2, Application US/08900953
; Patent No. 5994096
; GENERAL INFORMATION:
; APPLICANT: Burnham,, Martin K.R.
; APPLICANT: Lonetto,, Michael A.
; APPLICANT: Warren,, Patrick V.
; TITLE OF INVENTION: No. 5994096el Regulator
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,953
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-900-953-2

Query Match 62.5%; Score 35; DB 2; Length 480;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9
Db 36 TGDWFYGM 43

RESULT 3
US-08-900-953-4
; Sequence 4, Application US/08900953
; Patent No. 5994096
; GENERAL INFORMATION:
; APPLICANT: Burnham,, Martin K.R.
; APPLICANT: Lonetto,, Michael A.
; APPLICANT: Warren,, Patrick V.
; TITLE OF INVENTION: No. 5994096el Regulator
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,953
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-900-953-4

Query Match 62.5%; Score 35; DB 2; Length 480;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9
Db 36 TGDWFYGM 43

RESULT 4
US-08-428-488-16
; Sequence 16, Application US/08428488
; Patent No. 5624894
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas S.
; TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE
; TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,488
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Mary Katherine
; REGISTRATION NUMBER: 26,254
; REFERENCE/DOCKET NUMBER: 028724-087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
```

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note= "Position 1 = p-Glu."  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /note= "Position 11 = Met-NH2."  
US-08-428-488-16

Query Match 60.7%; Score 34; DB 1; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.9;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
Db 5 NKFYGLM 11

RESULT 5  
US-08-796-598-7  
; Sequence 7, Application US/08796598  
; Patent No. 5827659  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; ADDRESSEE: Thibeault  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/796,598  
; FILING DATE: 07-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446,055  
; FILING DATE: 19-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLYNN Esq., Kerry A.  
; REGISTRATION NUMBER: 33,693  
; REFERENCE/DOCKET NUMBER: SYP-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-796-598-7

Query Match 60.7%; Score 34; DB 2; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.9;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 DKFYGLM 10  
Db 5 NKFYGLM 11

RESULT 6  
US-08-447-175A-7  
; Sequence 7, Application US/08447175A  
; Patent No. 5869240  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; ADDRESSEE: Thibeault, LLP  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,175A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 422  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAUSCHENBACH, Kurt  
; REGISTRATION NUMBER: 40,137  
; REFERENCE/DOCKET NUMBER: SYP-114  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-447-175A-7

Query Match 60.7%; Score 34; DB 2; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.9;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
Db 5 NKFYGLM 11

RESULT 7  
US-09-214-614-1  
; Sequence 1, Application US/09214614  
; Patent No. 6225100  
; GENERAL INFORMATION:  
; APPLICANT: Grund, Alan D.  
; APPLICANT: Maurina-Brunker, Julie  
; TITLE OF INVENTION: NOVEL ARYL SULFOTRANSFERASE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver

; STATE: Colorado  
; COUNTRY: U.S.A  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/09/214,614  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 3161-15-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-214-614-1

Query Match 60.7%; Score 34; DB 4; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.9;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
; :|||||  
Db 5 NKFYGLM 11

RESULT 8  
PCT-US91-08177-19  
; Sequence 19, Application PC/TUS9108177  
; GENERAL INFORMATION:  
; APPLICANT: Samal, Siba K  
; TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
; STREET: 1201 New York Avenue N.W., suite 1000  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/08177  
; FILING DATE: 19911104  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/608,937  
; FILING DATE: 05-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hight, David W  
; REGISTRATION NUMBER: 30,265  
; REFERENCE/DOCKET NUMBER: 20509-96711  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-962-4854  
; TELEFAX: 202-962-8300  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 572 amino acids

; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Bovine respiratory syncytial virus  
; STRAIN: FS-1  
; PCT-US91-08177-19  
  
Query Match 60.7%; Score 34; DB 5; Length 572;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 NTGDKFYGLM 10  
; :|||||  
Db 120 NSTKRFYGLM 129  
  
RESULT 9  
US-08-411-098-14  
; Sequence 14, Application US/08411098  
; Patent No. 5830755  
; GENERAL INFORMATION:  
; APPLICANT: HWU, PATRICK; NISHIMURA,  
; APPLICANT: MICHAEL; ROSENBERG, STEVEN A.  
; TITLE OF INVENTION: T-CELL RECEPTORS AND  
; TITLE OF INVENTION: THEIR USE IN THERAPEUTIC AND DIAGNOSTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/411,098  
; FILING DATE: 27-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAROL M. GRUPPI  
; REGISTRATION NUMBER: 37,341  
; REFERENCE/DOCKET NUMBER: 2026-4188  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32  
; TYPE: AMINO ACID  
; STRANDEDNESS: UNKNOWN  
; TOPOLOGY: UNKNOWN  
; US-08-411-098-14

Query Match 57.1%; Score 32; DB 2; Length 32;  
Best Local Similarity 71.4%; Pred. No. 19;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFY 7  
; :|||||  
Db 9 NTGNQFY 15

RESULT 10

US-09-154-802-3  
; Sequence 3, Application US/09154802  
; Patent No. 5989822  
; GENERAL INFORMATION:  
; APPLICANT: Y. Tom Tang  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: ATP SYNTHASE SUBUNIT HOMOLOG  
; FILE REFERENCE: PF-0596 US  
; CURRENT APPLICATION NUMBER: US/09/154,802  
; CURRENT FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: g1006572  
US-09-154-802-3

Query Match 57.1%; Score 32; DB 2; Length 255;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 NTGDKFYGL 9  
| | | | |  
Db 142 NHGAKFFGL 150

RESULT 11  
US-09-373-029-3  
; Sequence 3, Application US/09373029  
; Patent No. 6036954  
; GENERAL INFORMATION:  
; APPLICANT: Y. Tom Tang  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: ATP SYNTHASE SUBUNIT HOMOLOG  
; FILE REFERENCE: PF-0596 US  
; CURRENT APPLICATION NUMBER: US/09/373,029  
; CURRENT FILING DATE: 1999-08-11  
; EARLIER APPLICATION NUMBER: 09/154,802  
; EARLIER FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: g1006572  
US-09-373-029-3

Query Match 57.1%; Score 32; DB 3; Length 255;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 NTGDKFYGL 9  
| | | | |  
Db 142 NHGAKFFGL 150

RESULT 12  
US-08-960-048-11  
; Sequence 11, Application US/08960048C  
; Patent No. 6271443  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; TITLE OF INVENTION: Sequences  
; FILE REFERENCE: 15621/01/US  
; CURRENT APPLICATION NUMBER: US/08/960,048C  
; CURRENT FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILING DATE: 1996-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 693  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-08-960-048-11

Query Match 57.1%; Score 32; DB 4; Length 693;  
Best Local Similarity 60.0%; Pred. No. 3.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
| | | | |  
Db 230 NEGTLFYGLV 239

RESULT 13  
US-08-894-483-8  
; Sequence 8, Application US/08894483  
; Patent No. 6054163  
; GENERAL INFORMATION:  
; APPLICANT: WETTENHALL, RICHARD E.H.  
; APPLICANT: DAVIDSON, BARRIE E.  
; APPLICANT: HILLIER, ALAN J.  
; APPLICANT: HARMARK, KIM  
; APPLICANT: JACK, RALPH W.  
; APPLICANT: HICKEY, MALCOLM W.  
; APPLICANT: COVENTRY, JOHN  
; APPLICANT: WAN, JASON  
; TITLE OF INVENTION: NOVEL BACTERIOICIN PISCICOLIN 126  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye, P.C.  
; STREET: 1100 No. 6054163th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,483  
; FILING DATE: 22-FEB-1996  
; CLASSIFICATION: 426  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PNI310  
; FILING DATE: 22-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mitchard, Leonard C.  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 47-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4005  
; TELEFAX: (703) 816-4100  
; TELEX: N/A  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 62 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear



; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Carnobacterium piscicola  
; STRAIN: JG126  
; US-08-894-483-8

Query Match 55.4%; Score 31; DB 3; Length 62;  
Best Local Similarity 71.4%; Pred. No. 54;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFG 8  
|||  
DB 16 TGGKYG 22

RESULT 14  
US-08-941-445A-7  
; Sequence 7, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-941-445A-7

Query Match 55.4%; Score 31; DB 3; Length 609;  
Best Local Similarity 71.4%; Pred. No. 5.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFG 8  
|||  
DB 178 TGEKIYG 184

RESULT 15

US-08-433-318A-67  
; Sequence 67, Application US/08433318A  
; Patent No. 5663144  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5663144rls  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: WORDPERFECT 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/433,318A  
; FILING DATE: 03-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark DeLuca  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: UPN-2106  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-433-318A-67

Query Match 53.6%; Score 30; DB 1; Length 6;  
Best Local Similarity 83.3%; Pred. No. 1.7e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GDKFG 8  
|||  
DB 1 GDAFYG 6

Search completed: June 19, 2002, 08:21:14  
Job time: 71 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:20:03 ; Search time 51.85 Seconds  
(without alignments)  
21.422 Million cell updates/sec

Title: US-09-713-299B-1  
Perfect score: 56  
Sequence: 1 NTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	67.9	93	22	AAU23760
2	38	67.9	867	22	Novel human enzyme
3	37	66.1	348	22	E. coli growth and
4	37	66.1	440	22	Putative P. abyssi
5	37	66.1	539	22	Drosophila melanog
6	37	66.1	572	13	Pseudomonas aerugi
7	36	64.3	280	21	Bovine RSV strain
8	36	64.3	282	21	Arabidopsis thalia
9	36	64.3	288	21	Arabidopsis thalia
10	36	64.3	290	21	Arabidopsis thalia
11	36	64.3	314	22	Corynebacterium gl

12	36	64.3	326	21	AAG24814
13	36	64.3	328	21	Arabidopsis thalia
14	36	64.3	334	22	C glutamicum prote
15	36	64.3	349	21	Arabidopsis thalia
16	36	64.3	349	21	Arabidopsis thalia
17	36	64.3	398	21	Arabidopsis thalia
18	36	64.3	398	21	Arabidopsis thalia
19	36	64.3	416	21	Arabidopsis thalia
20	36	64.3	416	21	Arabidopsis thalia
21	35	62.5	155	20	M. leprae RP-facto
22	35	62.5	220	20	M. luteus RP-facto
23	35	62.5	220	20	M. luteus RP-facto
24	35	62.5	274	22	Drosophila melanog
25	35	62.5	480	20	Staphylococcus aur
26	35	62.5	480	20	Staphylococcus aur
27	35	62.5	531	22	Novel human diagno
28	35	62.5	1695	22	Amino acid sequenc
29	34	60.7	11	18	Physalaemin peptid
30	34	60.7	11	19	Tyrosylpeptide phy
31	34	60.7	11	22	Trachykinins peptid
32	34	60.7	12	6	Hylambatin dodecap
33	34	60.7	138	22	S. epidermidis ope
34	34	60.7	341	22	Drosophila melanog
35	34	60.7	509	22	Drosophila melanog
36	34	60.7	567	13	Bovine RSV strain
37	34	60.7	574	13	Sequence of protei
38	34	60.7	1278	22	Drosophila melanog
39	34	60.7	1560	10	S. cremoris protei
40	34	60.7	1959	12	Mutant protease (d
41	34	60.7	1962	12	Mutant protease (K
42	34	60.7	1962	12	Mutant protease (N
43	34	60.7	1962	12	Mutant protease (A
44	34	60.7	1962	12	Mutant protease (A
45	34	60.7	1962	12	Mutant protease (A

ALIGNMENTS

RESULT 1

AAU23760  
ID AAU23760 standard; Protein; 93 AA.  
XX AC AAU23760;  
XX DT 17-DEC-2001 (first entry)  
XX DE Novel human enzyme polypeptide #846.

Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
ligase; hyperproliferative disorder; immunodeficiency disorder;  
autoimmune disorder; neurological disorder; metabolic disorder;  
inflammatory disorder; cardiovascular disorder; reproductive disorder;  
blood-related disorder; infectious disorder; cytostatic; anti arthritic;  
nephrotropic; anticoagulant.

OS Homo sapiens.

PN WO200155301-A2.

PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01239.

XX PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218280.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465566/50.  
XX N-PSDB; AAS41630.  
PT Novel polypeptides and polynucleotides useful for diagnosing,  
PT preventing, treating neural, immune system, muscular, reproductive,  
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
PT diseases -  
XX Claim 11; SEQ ID No 1756; 1180pp; English.  
PS The present invention relates to the isolation of novel human enzyme  
XX polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,

CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
 CC blood-related disorders (e.g. haemophilia), reproductive disorders  
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The  
 CC polynucleotides of the invention can also be used in gene therapy.  
 CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the  
 CC invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 93 AA;

Query Match 67.9%; Score 38; DB 22; Length 93;  
 Best Local Similarity 87.5%; Pred. NO. 14;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GDKFYGLM 10  
 | |||||  
 Db 18 gkkyfgilm 25

RESULT 2  
 AAG98887  
 ID AAG98887 standard; Protein; 867 AA.

XX AAG98887;

DT 26-SEP-2001 (first entry)

XX E. coli growth and proliferation related protein sequence SEQ ID NO:357.

XX Escherichia coli; growth; proliferation; microbial; antimicrobial;  
 KW bacterial infection; microorganism.

XX Escherichia coli.

PN WO200134810-A2.

PD 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30950.

PR 09-NOV-1999; 99US-0164415.

XX (ELIT-) ELITRA PHARM INC.

PI Forsyth RA, Ohlsen K, Zyskind J;

XX WPI; 2001-335933/35.

DR N-PSDB; AAH84558.

PT Novel nucleic acids that inhibit Escherichia coli proliferation, useful  
 PT for screening for homologous genes and for designing expression vectors

PS Claim 19; Page 432-434; 522pp; English.

XX AAH84373 to AAH84499 represent Escherichia coli growth and proliferation  
 CC related DNA sequences (I). AAH84500 to AAH84670 encode the E. coli  
 CC growth and proliferation related proteins given in AAG99078 and AAG98830  
 CC to AAG98999. (I) can be used as potential targets for the generation of  
 CC new antimicrobial agents, and for identification of compounds which  
 CC interact with the gene products of (I). In addition the expression of  
 CC (I) and the purification of the proteins, the purified proteins can be  
 CC used to generate reagents and screen small molecule libraries or other  
 CC candidate compound libraries for compounds that can be further developed

CC to yield novel antimicrobial compounds. In addition, nucleic acid probes  
 CC complementary to (I) that are specific for particular species of  
 CC microorganisms can be used to identify particular microorganism species  
 CC in clinical specimens, therefore, providing a rapid and dependable  
 CC method by which to identify the causative agents of a bacterial  
 CC infection. Also, antibodies generated against proteins translated from  
 CC mRNA transcribed from proliferation-required sequences can also be used  
 CC to screen for specific microorganisms that produce such proteins in a  
 CC species-specific manner. AAH84371 and AAH84670 represent sequencing  
 CC primers used in the isolation of E. coli growth and proliferation  
 CC related sequence, which are used in an example from the present  
 CC invention.

XX Sequence 867 AA;

Query Match 67.9%; Score 38; DB 22; Length 867;  
 Best Local Similarity 80.0%; Pred. NO. 1.4e+02;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10  
 ||||| |||  
 Db 92 ntgdksggilm 101

RESULT 3  
 AAB96784  
 ID AAB96784 standard; Protein; 348 AA.

XX AAB96784;

DT 29-OCT-2001 (first entry)

DE Putative P. abyssi threonine dehydratase/threonine synthase #2.

XX Hyperthermophilic archaeon; hyperthermophilic protein.

OS Pyrococcus abyssi.

PN FR2792651-A1.

XX 27-OCT-2000.

XX 21-APR-1999; 99FR-0005034.

PR 21-APR-1999; 99FR-0005034.

XX (CNRS ) CNRS CENT NAT RECH SCI.  
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.

XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
 XX Querellou J, Weissenbach J, Saurin W, Heilig R;

XX WPI; 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssi encode  
 XX proteins useful in industry -

PS Claim 7; Pages 1555-1556; 1657pp; French.

XX The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins P. abyssi is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such P. abyssi protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO2000065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.

XX Sequence 348 AA;

Query Match 66.1%; Score 37; DB 22; Length 348;  
Best Local Similarity 77.8%; Pred. No. 81;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10  
|||  
Db 305 tgfafyglm 313

## RESULT 4

ABB72014  
ID ABB72014 standard; Protein; 440 AA.

XX AC ABB72014;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 42834.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL16117.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX PS Disclosure; SEQ ID NO 42834; 2lpp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 440 AA;

Query Match 66.1%; Score 37; DB 22; Length 440;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTGDKFY 7

Db 137 ntgdefy 143

## RESULT 5

AQU33643

ID AU33643 standard; Protein; 539 AA.

XX AC AU33643;

XX DT 14-FEB-2002 (first entry)

XX DE Pseudomonas aeruginosa cellular proliferation protein #87.

XX KW Antisense; prokaryotic cellular proliferation protein;  
KW antibiotic; antibacterial; drug design.

XX OS Pseudomonas aeruginosa.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207727P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR N-PSDB; AAS51502.

XX PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 5139; 51lpp; English.

XX CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 539 AA;

Query Match 66.1%; Score 37; DB 22; Length 539;  
Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9

Db 504 tgdifygl 511

## RESULT 6

```

AAR24189
ID  AAR24189 standard; Protein; 572 AA.
XX
XX  AAR24189;
AC
XX  24-NOV-1992 (first entry)
DT
XX
XX  Bovine RSV strain A 51908 F protein.
DE
XX
XX  Bovine respiratory syncytial virus; vaccine; diagnosis; antibodies;
KW  F gene; fusion protein; BRSV.
KW
XX
XX  Bovine respiratory syncytial virus strain A 51908.
OS
XX
XX  WO9207940-A.
PN
XX
XX  14-MAY-1992.
PD
XX
XX  04-NOV-1991; 91WO-US08177.
PF
XX
XX  05-NOV-1990; 90US-0608937.
PR
XX
XX  (SAMA/) SAMAL S K.
PA
XX
XX  Samal SK;
PI
XX
XX  WPI; 1992-183675/22.
DR
XX
XX  N-PSDB; AAQ25033.
DR
XX
XX  Bovine respiratory syncytial virus genes - used in the prodn. of
PT  agents for use in detection and as vaccines for BRSV infection.
PT
XX
XX  Claim 30; Page 50; 74pp; English.
XX
XX  This is the sequence of bovine respiratory syncytial virus (BRSV)
CC  strain AA51908 fusion (F) protein. The F protein has a molecular
CC  weight of 66-68 kD and has a rod shaped morphology, suggesting that
CC  it may be part of the studied projections of the virus. It is an
CC  envelope protein and as such is essential for a recombinant vaccine.
CC  It comprises of two smaller glycoproteins linked by disulphide
CC  bonds. One of them has a molecular weight of about 43-56 kD and the
CC  other between about 19 and 22kD. The F protein has been show to be
CC  the fusion protein by inhibition of cell fusion by a monoclonal
CC  antibody to the F protein. It can be used in the detection of BRSV
CC  antibodies and in vaccines to prevent infection. It can also be used
CC  for the production of BRSV protein antibodies.
CC  See also AAR24184-R24191 and AAR25310.
XX
XX  Sequence 572 AA;
SQ

Query Match 66.1%; Score 37; DB 13; Length 572;
Best Local Similarity 70.0%; Pred No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTGDKFYGLM 10
   I: |||||
Db 120 nstkkfyglm 129

RESULT 7
AAG24816
ID  AAG24816 standard; Protein; 280 AA.
XX
XX  AAG24816;
AC
XX
XX  17-OCT-2000 (first entry)
DT
XX
XX  Arabidopsis thaliana protein fragment SEQ ID NO: 28633.
DE
XX
XX  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.

```

```

XX  Arabidopsis thaliana.
OS
XX  EP1033405-A2.
PN
XX
XX  06-SEP-2000.
PD
XX
XX  25-FEB-2000; 2000EP-0301439.
PF
XX
XX  25-FEB-1999; 99US-0121825.
XX  05-MAR-1999; 99US-0123180.
XX  09-MAR-1999; 99US-0123548.
XX  23-MAR-1999; 99US-0125788.
XX  25-MAR-1999; 99US-0126264.
XX  29-MAR-1999; 99US-0126785.
XX  01-APR-1999; 99US-0127462.
XX  06-APR-1999; 99US-0128234.
XX  08-APR-1999; 99US-0128714.
XX  16-APR-1999; 99US-0129845.
XX  19-APR-1999; 99US-0130077.
XX  21-APR-1999; 99US-0130449.
XX  23-APR-1999; 99US-0130510.
XX  28-APR-1999; 99US-0130891.
XX  30-APR-1999; 99US-0131449.
XX  30-APR-1999; 99US-0132048.
XX  04-MAY-1999; 99US-0132407.
XX  05-MAY-1999; 99US-0132484.
XX  06-MAY-1999; 99US-0132485.
XX  06-MAY-1999; 99US-0132486.
XX  07-MAY-1999; 99US-0132487.
XX  11-MAY-1999; 99US-0132863.
XX  14-MAY-1999; 99US-0134256.
XX  14-MAY-1999; 99US-0134218.
XX  14-MAY-1999; 99US-0134219.
XX  14-MAY-1999; 99US-0134221.
XX  18-MAY-1999; 99US-0134370.
XX  19-MAY-1999; 99US-0134768.
XX  20-MAY-1999; 99US-0134941.
XX  21-MAY-1999; 99US-0135124.
XX  21-MAY-1999; 99US-0135353.
XX  24-MAY-1999; 99US-0135629.
XX  25-MAY-1999; 99US-0136021.
XX  27-MAY-1999; 99US-0136392.
XX  28-MAY-1999; 99US-0136782.
XX  01-JUN-1999; 99US-0137222.
XX  03-JUN-1999; 99US-0137528.
XX  04-JUN-1999; 99US-0137502.
XX  07-JUN-1999; 99US-0137724.
XX  08-JUN-1999; 99US-0138094.
XX  10-JUN-1999; 99US-0138540.
XX  10-JUN-1999; 99US-0138847.
XX  14-JUN-1999; 99US-0139119.
XX  16-JUN-1999; 99US-0139452.
XX  16-JUN-1999; 99US-0139453.
XX  17-JUN-1999; 99US-0139492.
XX  18-JUN-1999; 99US-0139454.
XX  18-JUN-1999; 99US-0139455.
XX  18-JUN-1999; 99US-0139456.
XX  18-JUN-1999; 99US-0139457.
XX  18-JUN-1999; 99US-0139458.
XX  18-JUN-1999; 99US-0139459.
XX  18-JUN-1999; 99US-0139460.
XX  18-JUN-1999; 99US-0139461.
XX  18-JUN-1999; 99US-0139462.
XX  18-JUN-1999; 99US-0139463.
XX  18-JUN-1999; 99US-0139750.
XX  18-JUN-1999; 99US-0139763.
XX  21-JUN-1999; 99US-0139817.
XX  22-JUN-1999; 99US-0139899.
XX  23-JUN-1999; 99US-0140353.
XX  23-JUN-1999; 99US-0140354.
XX  24-JUN-1999; 99US-0140695.
XX  28-JUN-1999; 99US-0140823.

```

PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 280;  
Best Local Similarity 75.0%; Pred. No. 99;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8  
||| | | |  
Db 90 ntgdyffg 97

RESULT 8  
AAG46430  
ID AAG46430 standard; Protein; 282 AA.  
XX  
AC AAG46430;

XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58412.

XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.

XX  
PD 06-SEP-2000.  
XX

PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145226.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.



```
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157733.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 64.38; Score 36; DB 21; Length 282;
Best Local Similarity 75.08; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8
Db 90 ntgdyffg 97
|||||

RESULT 9
AAG24815
ID AAG24815 standard; Protein; 288 AA.
AC AAG24815;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 28632.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
```

PR 19-JUL-1999; 99US-01443325.  
PR 19-JUL-1999; 99US-01443331.  
PR 19-JUL-1999; 99US-01443332.  
PR 19-JUL-1999; 99US-01443333.  
PR 19-JUL-1999; 99US-01443334.  
PR 19-JUL-1999; 99US-01443335.  
PR 20-JUL-1999; 99US-01443352.  
PR 20-JUL-1999; 99US-01443352.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.

PR \*12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 288;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8  
||| |:  
Db 98 ntgdyffg 105

RESULT 10  
AAG46429  
ID AAG46429 standard; Protein; 290 AA.  
XX AC AAG46429;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 58411.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 08-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144353.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 28-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.

```
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 290;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8
Db 98 ntgdyffg 105

RESULT 11
AAB79838
ID AAB79838 standard; Protein; 314 AA.
XX
AC AAB79838;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:410.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.

PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
XX (BADI ) BASF AG.
XX
PI Pompejus M, Kroege B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF71957.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 758-759; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 314 AA;

Query Match 64.3%; Score 36; DB 22; Length 314;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDRFYGLM 10
||||:|:
```

Db 159 tgdkyfgfl 167

RESULT 12  
AAG24814  
ID AAG24814 standard; Protein; 326 AA.  
XX AC AAG24814;  
XX AC AAG24814;  
DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28631.  
XX DE Arabidopsis thaliana.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
PD 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0158458.  
PR 29-SEP-1999; 99US-0158596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 326;

Best Local Similarity 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTGDKFYG 8

Db 136 ntgdyfig 143

RESULT 13

AAG46428

ID AAG46428 standard; Protein; 328 AA.

XX

AC -AAG46428;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58410.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132883.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.

```
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140981.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
```

```
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 64.3%; Score 36; DB 21; Length 328;

Best Local Similarity 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTGDKFYG 8

Db 136 ntgdyffg 143

RESULT 14

AAG91391

ID AAG91391 standard; Protein; 334 AA.

XX AC AAG91391;

XX AC AAG91391;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 5145.

XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.  
XX Corynebacterium glutamicum.  
XX EP1108790-A2.  
XX 20-JUN-2001.  
PD 18-DEC-2000; 2000EP-0127668.  
XX 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280988.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.  
DR N-PSDB; AAH66610.  
XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analysing  
PT expression profile or pattern of a gene and identifying homologous gene  
PT  
PS Claim 17; SEQ ID NO: 5145; 246pp + Sequence Listing; English.  
XX The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of coryneform bacterium, measuring expression amount and  
CC analysing the expression profile or expression pattern of a gene derived  
CC from Coryneform bacterium, and identifying a homologue of a gene derived  
CC from coryneform bacterium. Coryneform bacteria are useful for producing  
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a protein described  
CC in the exemplification of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.  
XX  
SQ Sequence 334 AA;  
  
Query Match 64.3%; Score 36; DB 22; Length 334;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 TGDKFYGLM 10  
||||:|  
Db 179 tgdkyfgfl 187  
  
RESULT 15  
AAG07312  
ID AAG07312 standard; Protein; 349 AA.  
XX  
AC AAG07312;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4418.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX

PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140931.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.



PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 349;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDKFYG 8

Db 130 gdkfyg 135

Search completed: June 19, 2002, 08:24:08  
Job time: 245 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:24:42 ; Search time 26.86 Seconds  
(without alignments)  
35.774 Million cell updates/sec

Title: US-09-713-299B-2  
Perfect score: 56  
Sequence: 1 DTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	100.0	10	2 B49581	sialokinin II - ye
2	51	91.1	10	2 A49581	sialokinin I - yel
3	41	73.2	2206	1 GNNY21	genome polyprotein
4	39	69.6	10	2 A24867	scyllorhinin I - s
5	39	69.6	2214	1 A48548	genome polyprotein
6	37	66.1	129	2 G98225	hypothetical prote
7	37	66.1	257	2 T47491	hypothetical prote
8	37	66.1	230	2 C75192	probable l-aminocy
9	37	66.1	539	2 E83181	probable MFS trans
10	37	66.1	979	2 T40006	hypothetical prote
11	37	66.1	1134	2 G69269	molybdopterin-bind
12	37	66.1	1157	2 F97255	fusion of alpha-gl
13	37	66.1	1217	2 F97177	alpha-glucosidase
14	36	64.3	11	2 A10333	ranatachykinin A -
15	36	64.3	308	2 A52829	hypothetical prote
16	36	64.3	333	2 T45023	hypothetical prote
17	36	64.3	337	2 C97607	hypothetical prote
18	36	64.3	381	2 B84935	acetylornithine de
19	36	64.3	416	2 D84638	probable sarcosine
20	36	64.3	649	2 G71283	probable ATP-depen
21	35	62.5	109	2 S74409	hypothetical prote
22	35	62.5	161	2 AD2540	hypothetical prote
23	35	62.5	213	2 S13398	hypothetical prote
24	35	62.5	264	2 S52643	chloramphenicol O-
25	35	62.5	267	2 H90368	phycobilisome matu
26	35	62.5	311	2 T15997	hypothetical prote
27	35	62.5	405	2 T41593	hypothetical prote
28	35	62.5	548	2 T47548	hypothetical prote
29	35	62.5	736	2 D86271	protein F16A14.2 [

RESULT 1  
B49581  
sialokinin II - yellow fever mosquito  
C;Species: Aedes aegypti (yellow fever mosquito)  
C;Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 18-Aug-2000  
C;Accession: B49581  
R;Champagne, D.E.; Ribeiro, J.M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 138-142, 1994  
A;Title: Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito  
A;Reference number: A49581; MUID:94105119  
A;Accession: B49581  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <CHA>  
A;Experimental source: Rockefeller, salivary gland  
A;Note: sequence extracted from NCBI backbone (NCBIP:141842)  
C;Superfamily: unassigned animal peptides

Query Match 100.0%; Score 56; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10  
| | | | | | | | | |  
Db 1 DTGDKFYGLM 10

RESULT 2  
A49581  
sialokinin I - yellow fever mosquito  
C;Species: Aedes aegypti (yellow fever mosquito)  
C;Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 18-Aug-2000  
C;Accession: A49581  
R;Champagne, D.E.; Ribeiro, J.M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 138-142, 1994  
A;Title: Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito  
A;Reference number: A49581; MUID:94105119  
A;Accession: A49581  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <CHA>  
A;Experimental source: Rockefeller, salivary gland  
A;Note: sequence extracted from NCBI backbone (NCBIP:141841)  
C;Superfamily: unassigned animal peptides

Query Match 91.1%; Score 51; DB 2; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.0019;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DTGDKFYGLM 10

arylphorin precurs  
peroxinectin precu  
probable athila tr  
ranatachykinin B -  
physalaemin - frog  
tachykinin - Afric  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
transposase ISC105  
hypothetical prote  
asparagine-rich bl  
two-component sens  
early E1B 54K prot  
aryl hydrocarbon (

ALIGNMENTS

Db 1 NTGDKFYGLM 10

### RESULT

GENOME polypeptide - coxsackievirus A21 (strain Coe)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein  
polymerase (EC 2.7.7.48)  
C:Species: coxsackievirus A21  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jun-2000  
C:Accession: A33373  
R:Hughes, P.J.; North, C.; Minor, P.D.; Stanway, G.  
J. Gen. Virol. 70, 2943-2952, 1989  
A:title: The complete nucleotide sequence of coxsackievirus A21.

Query Match 73.2%; Score 41; DB 1; Length 2206;  
Best Local Similarity 70.0%; Pred. No. 36;  
Matches 7; Conservative 1; Mismatches 2; Indels

```
QY      1 DTGDKFYGLM 10
        | | | | | | :
Db      799 DAGDTFYGLV 808
```

## RESULTS

scyllorhinin I - smaller spotted catshark  
C:Species: Scyllorhinus canicula (smaller spotted dogfish)  
C:Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 18-Aug-2000  
C:Accession: A24867; S3301  
R:Conlon, J.M.; Deacon, C.F.; O'Toole, L.; Thim, L.  
FEBS Lett. 200, 111-116, 1986  
A:Title: Scyllorhinin I and II: two novel tachykinins from dogfish gut.

Query Match 69.6%; Score 39; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.34;

```
Matches      7: Conservative      0: Mismatches      0: Indels      0: Gaps      0:
```

QY 4 DKFYGLM 10  
11111111

Db 4 DKFYGLM 10

## 5 RESULT

genome polypeptide - coxsackievirus A24 (strain EH24/70)  
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core  
 polymerase (EC 2.7.7.48)  
 C:Species: coxsackievirus A24  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 05-Jun-1998  
 C:Accession: A48548  
 R:Supanaranond, K.; Takeda, N.; Yamazaki, S.  
 Virus Genes 6, 149-158, 1992  
 A:Title: The complete nucleotide sequence of a variant of Coxsackievirus A24, an agent  
 A:Reference number: A48548; MUID:92271460

Query Match 69.6%; Score 39; DB 1; Length 2214;  
Best Local Similarity 60.0%; Pred. No. 86;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGLM 10  
|:|:|:|:|:  
Db 805 DSGDTYYGLV 814

RESULTS

hypothetical protein AGR\_L1500 [imported] - Agrobacterium tumefaciens (strain C58, C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: G98225  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldmann, S.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium A:Reference number: A97359; PMID:11743194

Query Match 66.1%; Score 37; DB 2; Length 129;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 2; Indels

QY 1 DTGDKFYGL 9  
 Db 14 DTGKRFYDL 22

RESULT 7  
 T47491  
 hypothetical protein F9K21.90 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47491  
 R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.;  
 submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24467  
 A:Accession: T47491  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-257 <JOP>  
 A:Cross-references: EMBL:AL138657  
 A:Experimental source: cultivar Columbia; BAC clone F9K21  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 213/3  
 A:Note: F9K21.90

Query Match 66.1%; Score 37; DB 2; Length 257;  
 Best Local Similarity 66.7%; Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
 Db 240 ETGGRFYGL 248

RESULT 8  
 C75192  
 probable 1-aminocyclopropane-1-carboxylate deaminase (EC 4.1.99.4) PAB2303 [similarity]  
 N:Alternate names: ACC deaminase  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: C75192  
 R:Anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
 A:Reference number: A75001  
 A:Accession: C75192  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-330 <KAW>  
 A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB48986.1; PID:g545749  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB2303  
 C:Function:  
 A:Description: catalyzes the irreversible conversion of 1-aminocyclopropane-1-carboxylat  
 C:Superfamily: 1-aminocyclopropane-1-carboxylate deaminase  
 C:Keywords: carbon-carbon lyase; phosphoprotein; pyridoxal phosphate  
 F:54/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 66.1%; Score 37; DB 2; Length 330;  
 Best Local Similarity 77.8%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10  
 Db 287 TGKAFYGLM 295

RESULT 9  
 E93181  
 probable MFS transporter PA3709 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: E83181  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: E83181  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-539 <STO>  
 A:Cross-references: GB:AE004790; GB:AE004091; NID:g9949873; PIDN:AAG07096.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA3709

Query Match 66.1%; Score 37; DB 2; Length 539;  
 Best Local Similarity 87.5%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9  
 Db 504 TGDIFYGL 511

RESULT 10  
 T40006  
 hypothetical protein SPBC25H2.11c - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Sep-2000  
 C:Accession: T40006  
 R:Durso, G.; Iye, G.; Bowman, S.; Church, C.; Wood, V.; Barrell, B.G.; Rajandream, M.  
 submitted to the EMBL Data Library, April 1997  
 A:Reference number: Z21898  
 A:Accession: T40006  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-979 <DUR>  
 A:Cross-references: EMBL:Z95397; PIDN:CAB08777.1; GSPDB:GN00067; SPDB:SPBC25H2.11c  
 A:Experimental source: strain 972h-; cosmid c25H2  
 C:Genetics:  
 A:Gene: SPDB:SPBC25H2.11c  
 A:Map position: 2  
 A:Introns: 192/1; 258/3; 292/3; 339/1; 753/3; 801/3; 873/2  
 C:Superfamily: bromodomain homology  
 F:324-379/Domain: bromodomain homology <BRO>

Query Match 66.1%; Score 37; DB 2; Length 979;  
 Best Local Similarity 66.7%; Pred. No. 89;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
 Db 830 ETGDDFFGL 838

RESULT 11  
 G69269  
 molybdopterin-binding oxidoreductase homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000  
 C:Accession: G69269  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes,  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343  
A:Accession: G69269  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1134 <KLE>  
A:Cross-references: GB:AE001095; GB:AE000782; NID:g2689418; PIDN:AA991069.1; PID:g265048

Query Match 66.1%; Score 37; DB 2; Length 1134;  
Best Local Similarity 85.7%; Pred. No. 1.e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGDGFYGL 8  
Db 890 TGEKFG 896  
|||||

RESULT 12  
F97255  
fusion of alpha-glucosidase (family 31 glycosyl hydrolase) and glycosidase (TreA/Mals fa  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: F97255  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: F97255  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1157 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80833.1; PID:g15025938; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2891

Query Match 66.1%; Score 37; DB 2; Length 1157;  
Best Local Similarity 55.6%; Pred. No. 1.e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
Db 165 NSGDRFYGI 173  
|||||

RESULT 13  
F97177  
alpha-glucosidase fused to unknown alpha-amylase C-terminal. domain [imported] - Clostri  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: F97177  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: F97177  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1217 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80209.1; PID:g15025253; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2252

Query Match 66.1%; Score 37; DB 2; Length 1217;  
Best Local Similarity 55.6%; Pred. No. 1.e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
|||||

Db 169 NSGDRFYGI 177  
|||||

RESULT 14  
A61033  
ranatachykinin A - bullfrog  
C:Species: Rana catesbeiana (bullfrog)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Aug-2000  
C:Accession: A61033; JE0426  
R:Kangawa, K.; Kozawa, H.; Hino, J.; Minamino, N.; Matsuo, H.  
Regul. Pept. 42(Suppl.1), S12, 1992  
A:Title: Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and i  
A:Reference number: A61033  
A:Accession: A61033  
A:Molecule type: protein  
A:Residues: 1-11 <KAN>  
R:Kozawa, H.; Hino, J.; Minamino, N.; Kangawa, K.; Matsuo, H.  
Biochem. Biophys. Res. Commun. 177, 588-595, 1991  
A:Title: Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and i  
A:Reference number: JE0426; MUID:91254337  
A:Accession: JE0426  
A:Molecule type: protein  
A:Residues: 1-11 <KOZ>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; neuropeptide  
F:11/Modified site: amidated carboxyl end (Met) #status experimental

Query Match 64.3%; Score 36; DB 2; Length 11;  
Best Local Similarity 85.7%; Pred. No. 1.4;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
Db 5 DREYGLM 11  
|||||

RESULT 15  
AE2829  
hypothetical protein Atu2060 [imported] - Agrobacterium tumefaciens (strain C58, Dupo  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AE2829  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AE2829  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AAL43051.1; PID:g17740518; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2060  
A:Map position: circular chromosome

Query Match 64.3%; Score 36; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDKFYG 8  
Db 212 GDKFYG 217  
|||||

Search completed: June 19, 2002, 08:24:43  
Job time: 260 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:21:14 ; Search time 21.94 Seconds  
(without alignments)  
11.133 Million cell updates/sec

Title: US-09-713-299B-2  
Perfect score: 56  
Sequence: 1 DTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	62.5	480	2	US-08-900-953-2
2	35	62.5	480	2	US-08-900-953-4
3	34	60.7	11	1	US-08-428-488-16
4	34	60.7	11	2	US-08-796-598-7
5	34	60.7	11	2	US-08-447-175A-7
6	34	60.7	11	4	US-09-214-614-1
7	34	60.7	2154	2	US-08-841-349-4
8	33	58.9	122	4	US-08-392-459-30
9	33	58.9	122	5	PCT-US91-08525-30
10	33	58.9	122	5	PCT-US93-04384-10
11	33	58.9	125	4	US-08-392-459-54
12	33	58.9	125	5	PCT-US91-08525-54
13	33	58.9	125	5	PCT-US93-04384-14
14	33	58.9	190	4	US-08-392-459-20
15	33	58.9	190	5	PCT-US91-08525-20
16	33	58.9	198	4	US-09-163-833-2
17	33	58.9	281	4	US-08-936-165A-323
18	33	58.9	572	5	PCT-US91-08177-11
19	33	58.9	748	4	US-08-392-459-24
20	33	58.9	748	4	US-08-392-459-28
21	33	58.9	748	4	US-08-392-459-34
22	33	58.9	748	5	PCT-US91-08525-24
23	33	58.9	748	5	PCT-US91-08525-28
24	33	58.9	748	5	PCT-US91-08525-34
25	33	58.9	748	5	PCT-US93-04384-4
26	33	58.9	748	5	PCT-US93-04384-6
27	33	58.9	1454	4	US-08-392-459-22

28	33	58.9	1454	4	US-08-392-459-26
29	33	58.9	1454	4	US-08-392-459-32
30	33	58.9	1454	5	PCT-US91-08525-22
31	33	58.9	1454	5	PCT-US91-08525-26
32	33	58.9	1454	5	PCT-US91-08525-32
33	33	58.9	1454	5	PCT-US93-04384-2
34	33	58.9	1454	5	PCT-US93-04384-8
35	33	58.9	1454	5	PCT-US93-04384-12
36	33	58.9	1454	5	PCT-US93-04384-16
37	33	58.9	1454	5	PCT-US93-04384-43
38	33	58.9	1454	5	PCT-US93-04384-44
39	33	58.9	1454	5	PCT-US93-04384-45
40	33	58.9	1454	5	PCT-US93-04384-46
41	33	58.9	1454	5	PCT-US93-04384-47
42	33	58.9	1454	5	PCT-US93-04384-48
43	31	55.4	62	3	US-08-894-483-8
44	31	55.4	114	2	US-08-222-719-5
45	31	55.4	114	2	US-08-470-925-5

ALIGNMENTS

RESULT 1  
US-08-900-953-2  
; Sequence 2, Application US/08900953  
; Patent No. 5994096  
; GENERAL INFORMATION:  
; APPLICANT: Burnham,, Martin K.R.  
; APPLICANT: Lonetto,, Michael A.  
; APPLICANT: Warren,, Patrick V.  
; TITLE OF INVENTION: No. 5994096el Regulator  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/900,953  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dickinson, Q. Todd  
; REGISTRATION NUMBER: 28,354  
; REFERENCE/DOCKET NUMBER: GMI0054  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215/994-2252  
; TELEFAX: 215/994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 480 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-900-953-2

Query Match 62.5%; Score 35; DB 2; Length 480;  
Best Local Similarity 75.0%; Pred. No. 82;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9  
 ||| |||  
 Db 36 TGDWFGM 43

## RESULT 2

US-08-900-953-4  
 ; Sequence 4, Application US/08900953  
 ; Patent No. 5994096  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burnham,, Martin K.R.  
 ; APPLICANT: Lonetto,, Michael A.  
 ; APPLICANT: Warren,, Patrick V.  
 ; TITLE OF INVENTION: No. 5994096el Regulator  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dechert, Price & Rhoads  
 ; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19103-2793  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/900,953  
 ; FILING DATE:  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dickinson, Q. Todd  
 ; REGISTRATION NUMBER: 28,354  
 ; REFERENCE/DOCKET NUMBER: GM10054  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215/994-2252  
 ; TELEFAX: 215/994-2222  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 480 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-900-953-4

Query Match 62.5%; Score 35; DB 2; Length 480;  
 Best Local Similarity 75.0%; Pred. No. 82;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGL 9  
 ||| |||  
 Db 36 TGDWFGM 43

## RESULT 3

US-08-428-488-16  
 ; Sequence 16, Application US/08428488  
 ; Patent No. 5624894  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BODOR, Nicholas S.  
 ; TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE  
 ; TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM  
 ; NUMBER OF SEQUENCES: 107  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis  
 ; STREET: P.O. Box 1404  
 ; CITY: Alexandria

STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/428,488  
 FILING DATE: 27-APR-1995  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Baumeister, Mary Katherine  
 REGISTRATION NUMBER: 26,254  
 REFERENCE/DOCKET NUMBER: 028724-087  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 11 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 1  
 OTHER INFORMATION: /note= "Position 1 = p-Glu."  
 FEATURE:  
 NAME/KEY: Modified-site  
 LOCATION: 11  
 OTHER INFORMATION: /note= "Position 11 = Met-NH2."  
 US-08-428-488-16

Query Match 60.7%; Score 34; DB 1; Length 11;  
 Best Local Similarity 85.7%; Pred. No. 2.7;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
 :|||||  
 Db 5 NKFYGLM 11

## RESULT 4

US-08-796-598-7  
 ; Sequence 7, Application US/08796598  
 ; Patent No. 5827659  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PATTERSON, DALE H.  
 ; APPLICANT: TARR, GEORGE E.  
 ; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
 ; TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
 ; ADDRESSEE: Thibeault  
 ; STREET: High Street Tower, 125 High Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/796,598  
 ; FILING DATE: 07-FEB-1997  
 ; CLASSIFICATION: 435



;; PRIOR APPLICATION DATA: 60.7%; Score 34; DB 2; Length 11;  
;; APPLICATION NUMBER: US 08/446,055  
;; FILING DATE: 19-MAY-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: FLYNN Esq., Kerry A.  
;; REGISTRATION NUMBER: 33,693  
;; REFERENCE/DOCKET NUMBER: SYP-115  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 248-7000  
;; TELEFAX: (617) 248-7100  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-796-598-7

Query Match 60.7%; Score 34; DB 2; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
:|||||  
Db 5 NKFYGLM 11

RESULT 5  
US-08-447-175A-7  
; Sequence 7, Application US/08447175A  
; Patent No. 5869240  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS  
; TITLE OF INVENTION: SPECTROMETRY.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; ADDRESSEE: Thibault, LLP  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,175A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 422  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAUSCHENBACH, Kurt  
; REGISTRATION NUMBER: 40,137  
; REFERENCE/DOCKET NUMBER: SYP-114  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-447-175A-7

Query Match 60.7%; Score 34; DB 2; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 DKFYGLM 10  
:|||||  
Db 5 NKFYGLM 11  
RESULT 6  
US-09-214-614-1  
; Sequence 1, Application US/09214614  
; Patent No. 6225100  
; GENERAL INFORMATION:  
; APPLICANT: Grund, Alan D.  
; APPLICANT: Maurina-Brunker, Julie  
; TITLE OF INVENTION: NOVEL ARYLSULFOTRANSFERASE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/214,614  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 3161-15-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-214-614-1

Query Match 60.7%; Score 34; DB 4; Length 11;  
Best Local Similarity 85.7%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DKFYGLM 10  
:|||||  
Db 5 NKFYGLM 11

RESULT 7  
US-08-841-349-4  
; Sequence 4, Application US/08841349B  
; Patent No. 5955594  
; GENERAL INFORMATION:  
; APPLICANT: MISHRA, LOPA  
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT.  
; FILE REFERENCE: XX/PO44700S0  
; CURRENT APPLICATION NUMBER: US/08/841,349B  
; .CURRENT FILING DATE: 1997-04-30  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 2154

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-841-349-4

Query Match          60.7%; Score 34; DB 2; Length 2154;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTGDKF 6
Db 1896 DTGDKF 1901

RESULT 8
US-08-392-459-30
; Sequence 30, Application US/08392459
; Patent No. 6280974
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham, Corporation
; TITLE OF INVENTION: Recombinant Feline Coronavirus S
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08525
; FILING DATE: 19911114
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/698,927
; FILING DATE: 13-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/613,066
; FILING DATE: 14-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: SBC 14532B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 270-5015
; TELEFAX: (215) 270-5090
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-08525-30

Query Match          58.9%; Score 33; DB 5; Length 122;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
Db 74 DNGTKIYGL 82

RESULT 10
PCT-US93-04384-10
; Sequence 10, Application PC/TUS9304384
; GENERAL INFORMATION:
; APPLICANT: Miller, Timothy J.
; APPLICANT: Klepfer, Sharon
; APPLICANT: Reed, Albert Paul
; APPLICANT: Jones, Elaine V.
; TITLE OF INVENTION: Compositions and Methods for Vaccination
; TITLE OF INVENTION: Against Coronaviruses
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation - Corporate
; ADDRESSEE: Patents
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Query Match          58.9%; Score 33; DB 4; Length 122;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9
Db 74 DNGTKIYGL 82

RESULT 9
PCT-US91-08525-30
; Sequence 30, Application PC/TUS9108525
```

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: PCT/US93/04384  
; FILING DATE: 19930507  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/882,171  
; FILING DATE: 08-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/698,927  
; FILING DATE: 13-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,066  
; FILING DATE: 14-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schreck, Patricia A.  
; REGISTRATION NUMBER: 33,777  
; REFERENCE/DOCKET NUMBER: SBC H85009-1  
; TELEPHONE: (215) 270-5015  
; TELEFAX: (215) 270-5090  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 122 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US93-04384-10

Query Match 58.9%; Score 33; DB 5; Length 122;  
Best Local Similarity 66.7%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
Db 74 DNGTKIYGL 82

RESULT 11  
US-08-392-459-54  
; Sequence 54, Application US/08392459  
; Patent No. 6280974  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham, Corporation  
; TITLE OF INVENTION: Recombinant Feline Coronavirus S  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-2799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/392,459  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/698,927  
; FILING DATE: 13-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,066  
; FILING DATE: 14-NOV-1990  
; ATTORNEY/AGENT INFORMATION:

; NAME: King, William T.  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: SBC 14532B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 270-5015  
; TELEFAX: (215) 270-5090  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 125 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-392-459-54

Query Match 58.9%; Score 33; DB 4; Length 125;  
Best Local Similarity 66.7%; Pred. No. 48;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
Db 77 DNGTKIYGL 85

RESULT 12  
PCT-US91-08525-54  
; Sequence 54, Application PC/TUS9108525  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham, Corporation  
; TITLE OF INVENTION: Recombinant Feline Coronavirus S  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-2799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/08525  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/698,927  
; FILING DATE: 13-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,066  
; FILING DATE: 14-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T.  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: SBC 14532B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 270-5015  
; TELEFAX: (215) 270-5090  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 125 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US91-08525-54

Query Match 58.9%; Score 33; DB 5; Length 125;  
Best Local Similarity 66.7%; Pred. No. 48;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY      1 DTGDKFYGL 9
      | | | | |
Db      77 DNGTKIYGL 85

RESULT 13
PCT-US93-04384-14
; Sequence 14, Application PC/TUS9304384
; GENERAL INFORMATION:
; APPLICANT: Miller, Timothy J.
; APPLICANT: Klepfer, Sharon
; APPLICANT: Reed, Albert Paul
; APPLICANT: Jones, Elaine V.
; TITLE OF INVENTION: Compositions and Methods for Vaccination
; TITLE OF INVENTION: Against Coronaviruses
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation - Corporate
; ADDRESSEE: Patents
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04384
; FILING DATE: 19930507
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/882,171
; FILING DATE: 08-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/698,927
; FILING DATE: 13-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/613,066
; FILING DATE: 14-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Schreck, Patricia A.
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: SBC H85009-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 270-5015
; TELEFAX: (215) 270-5090
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-04384-14

Query Match      58.98; Score 33; DB 5; Length 125;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DTGDKFYGL 9
      | | | | |
Db      77 DNGTKIYGL 85

RESULT 14
US-08-392-459-20
; Sequence 20, Application US/08392459
; Patent No. 6280974
; GENERAL INFORMATION:

```

```

; APPLICANT: SmithKline Beecham, Corporation
; TITLE OF INVENTION: Recombinant Feline Coronavirus S
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,459
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/698,927
; FILING DATE: 13-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/613,066
; FILING DATE: 14-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: SBC 14532B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 270-5015
; TELEFAX: (215) 270-5090
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 190 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-392-459-20

Query Match      58.98; Score 33; DB 4; Length 190;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DTGDKFYGL 9
      | | | | |
Db      140 DNGTKIYGL 148

RESULT 15
PCT-US91-08525-20
; Sequence 20, Application PC/TUS9108525
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham, Corporation
; TITLE OF INVENTION: Recombinant Feline Coronavirus S
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08525

```

; FILING DATE: 19911114.  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/698,927  
; FILING DATE: 13-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/613,066  
; FILING DATE: 14-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T.  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: SBC 14532B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 270-5015  
; TELEFAX: (215) 270-5090  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 190 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US91-08525-20

Query Match 58.9%; Score=33; DB 5; Length 190;  
Best Local Similarity 66.7%; Pred. No. 74;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9  
| | | | |  
Db 140 DNGTKIYGL 148

Search completed: June 19, 2002, 08:21:15  
Job time: 72 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 08:24:08 ; Search time 51.85 seconds  
(without alignments)  
21.422 Million cell updates/sec

Title: US-09-713-299B-2  
Perfect score: 56  
Sequence: 1 DTGDKFYGLM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	67.9	93	22	AAU23760
2	37	66.1	282	22	AAU43435
3	37	66.1	348	22	AAU96784
4	37	66.1	539	22	AAU33643
5	36	64.3	314	22	AAU79838
6	36	64.3	334	22	AAU91391
7	36	64.3	349	21	AAU07312
8	36	64.3	349	21	AAU42891
9	36	64.3	398	21	AAU07311
10	36	64.3	398	21	AAU42890
11	36	64.3	416	21	AAU07310

12	36	64.3	416	21	AAU42889	Arabidopsis thalia
13	35	62.5	15	21	AAU29938	Human CTR4-4 scaff
14	35	62.5	274	22	ABB70933	Drosophila melanog
15	35	62.5	480	20	AAU04108	Staphylococcus aur
16	35	62.5	480	20	AAU04109	Staphylococcus aur
17	35	62.5	759	21	AAU51003	C. vicina arylphor
18	35	62.5	767	22	ABG06582	Novel human diagno
19	35	62.5	1295	22	ABG06581	Novel human diagno
20	35	62.5	1695	22	AAU83975	Amino acid sequenc
21	34	60.7	11	18	AAU04613	Phyalaemin peptid
22	34	60.7	11	19	AAU48280	Tyrosylpeptide phy
23	34	60.7	11	19	AAU91386	tachykinins peptid
24	34	60.7	12	6	AAU50357	Hylambatin dodecap
25	34	60.7	138	22	AAU82111	S. epidermidis ope
26	34	60.7	434	21	AAU45559	Arabidopsis thalia
27	34	60.7	488	21	AAU45558	Arabidopsis thalia
28	34	60.7	509	22	AAU61698	Drosophila melanog
29	34	60.7	535	21	AAU45557	Arabidopsis thalia
30	34	60.7	1120	20	AAU81642	Mouse elf protein.
31	34	60.7	1242	21	AAU53371	Human colon cancer
32	34	60.7	1278	22	ABB71882	Drosophila melanog
33	34	60.7	1529	21	AAU96744	A. terreus ORF1 es
34	34	60.7	2154	20	AAU81639	Mouse elf-1 protei
35	33	58.9	74	22	AAU04854	Human SGP014 phosp
36	33	58.9	77	22	AAU04110	Novel human diagno
37	33	58.9	122	13	AAU24399	Prod. of the S gen
38	33	58.9	125	13	AAU24512	Prod. of UCD-2 of
39	33	58.9	125	14	AAU42469	Feline infectious
40	33	58.9	138	21	AAU11837	Arabidopsis thalia
41	33	58.9	173	22	AAU41763	Human polypeptide
42	33	58.9	173	22	AAU41764	Human polypeptide
43	33	58.9	190	13	AAU24394	FIPV fusion protei
44	33	58.9	198	22	AAU73214	Murine phosphatase
45	33	58.9	198	22	AAU73219	Human phosphatase

ALIGNMENTS

RESULT 1

AAU23760	AAU23760 standard; Protein; 93 AA.
ID	AAU23760 standard; Protein; 93 AA.
XX	
AC	AAU23760;
XX	
DT	17-DEC-2001 (first entry)
XX	
DE	Novel human enzyme polypeptide #846.
XX	
KW	Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;
KW	autolysosome disorder; neurological disorder; metabolic disorder;
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW	nephrotropic; anticoagulant.
OS	Homo sapiens.
XX	
FN	WO200155301-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01239.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229533.  
PR 08-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234223.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PR (HUMA-) HUMAN GENOME SCI INC.  
PR Rosen CA, Barash SC, Ruben SM;  
WPI: 2001-465566/50.  
N-PSDB; AAS41630.  
Novel polypeptides and polynucleotides useful for diagnosing,  
preventing, treating neural, immune system, muscular, reproductive,  
pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
diseases -  
Claim 11; SEQ ID No 1756; 1180pp; English.  
The present invention relates to the isolation of novel human enzyme  
polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
encoding them. The enzyme polypeptides of the invention may comprise the  
functional classes of oxidoreductases, transferases, hydrolases, lyases,

CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
 CC blood-related disorders (e.g. haemophilia), reproductive disorders  
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The  
 CC polynucleotides of the invention can also be used in gene therapy.  
 CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the  
 CC invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 93 AA;

Query Match 67.9%; Score 38; DB 22; Length 93;  
 Best Local Similarity 87.5%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GDKFYGLM 10  
 | | | | |  
 Db 18 gkkfyglm 25

RESULT 2

AAU43435  
 ID AAU43435 standard; Protein; 282 AA.

XX AC AAU43435;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #4331.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO2000181581-A2.

XX 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208941P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-FSDB; AAS59520.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -

PS Example 1; SEQ ID No 4630; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 282 AA;

Query Match 66.1%; Score 37; DB 22; Length 282;  
 Best Local Similarity 66.7%; Pred. No. 62;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9

| | | | |  
 Db 272 dagdrllygl 280

RESULT 3

AAB96784

ID AAB96784 standard; Protein; 348 AA.

XX AC AAB96784;

XX 29-OCT-2001 (first entry)

DE Putative P. abyssii threonine dehydratase/threonine synthase #2.

KW Hyperthermophilic archaeon; hyperthermophilic protein.

OS Pyrococcus abyssii.

XX FR2792651-A1.

XX 27-OCT-2000.

XX 21-APR-1999; 99FR-0005034.

XX 21-APR-1999; 99FR-0005034.

XX (CNRS ) CNRS CENT NAT RECH SCI.  
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.

PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

PI Querellou J, Weissenbach J, Saurin W, Heilig R;

XX WPI; 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssii encode  
 PT proteins useful in industry -

PS Claim 7: Pages 1555-1556; 1657pp; French.

XX The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssii (see AAB96431 and AAB41223-7) and P. abyssii proteins. P. abyssii is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present invention is one such P. abyssii protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO2000065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,



CC AAH75903-AAH75920 and AAG66436.  
XX  
SQ Sequence 348 AA;

Query Match 66.1%; Score 37; DB 22; Length 348;  
Best Local Similarity 77.8%; Pred. No. 78;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGDIFYGLM 10  
||| |||||  
Db 305 tgkafyglm 313

## RESULT 4

AAU33643  
ID AAU33643 standard; Protein; 539 AA.

XX AC AAU33643;

XX DT 14-FEB-2002 (first entry)

XX DE Pseudomonas aeruginosa cellular proliferation protein #87.

XX KW Antisense; prokaryotic cellular proliferation protein;

XX KW antibiotic; antibacterial; drug design.

XX OS Pseudomonas aeruginosa.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207272P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR N-PSDB; AAS51502.

XX PT New polynucleotides for the identification and development of

XX PS antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 5139; 51lpp; English.

XX CC The invention relates to antisense inhibitors of genes essential to  
XX CC prokaryotic cellular proliferation, their use in identifying the  
XX CC genes, their use in the discovery of novel antibiotics, the essential  
XX CC genes themselves and the encoded proteins. The prokaryotes used are  
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX CC invention is also useful for the identification of potential new targets  
XX CC for antibiotic development. The antisense nucleic acids can also be used  
XX CC to identify proteins used in proliferation, to express these proteins,  
XX CC and to obtain antibodies capable of binding to the expressed proteins.  
XX CC The proteins can be used to screen compounds in rational drug discovery  
XX CC programmes. The antisense nucleic acid sequence is also useful to screen  
XX CC for homologous nucleic acids which are required for cell proliferation in  
XX CC a wide variety of organisms. The present sequence represents an  
XX CC essential prokaryotic cellular proliferation protein.  
XX CC Note: The sequence data for this patent did not form part  
XX CC of the printed specification, but was obtained in electronic

CC ' format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 539 AA;

Query Match 66.1%; Score 37; DB 22; Length 539;  
Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDIFYGL 9

||| |||||  
Db 504 tgdifygl 511

## RESULT 5

AAAB79838

ID AAAB79838 standard; Protein; 314 AA.

XX AC AAAB79838;

XX DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:410.

XX KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
XX KW fine chemical production; microorganism; organic acid; nucleoside;  
XX KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
XX KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
XX KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.

XX OS Corynebacterium glutamicum.

XX PN WO200100843-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB00923.

XX PR 25-JUN-1999; 99US-0141031.

XX PR 01-JUL-1999; 99DE-1030476.

XX PR 02-JUL-1999; 99US-0142101.

XX PR 08-JUL-1999; 99DE-1031415.

XX PR 08-JUL-1999; 99DE-1031418.

XX PR 08-JUL-1999; 99DE-1031419.

XX PR 08-JUL-1999; 99DE-1031420.

XX PR 08-JUL-1999; 99DE-1031424.

XX PR 08-JUL-1999; 99DE-1031428.

XX PR 08-JUL-1999; 99DE-1031434.

XX PR 08-JUL-1999; 99DE-1031435.

XX PR 08-JUL-1999; 99DE-1031443.

XX PR 08-JUL-1999; 99DE-1031453.

XX PR 08-JUL-1999; 99DE-1031457.

XX PR 08-JUL-1999; 99DE-1031465.

XX PR 08-JUL-1999; 99DE-1031478.

XX PR 08-JUL-1999; 99DE-1031510.

XX PR 08-JUL-1999; 99DE-1031541.

XX PR 08-JUL-1999; 99DE-1031573.

XX PR 08-JUL-1999; 99DE-1031592.

XX PR 08-JUL-1999; 99DE-1031632.

XX PR 08-JUL-1999; 99DE-1031634.

XX PR 08-JUL-1999; 99DE-1031636.

XX PR 09-JUL-1999; 99DE-1032125.

XX PR 09-JUL-1999; 99DE-1032126.

XX PR 09-JUL-1999; 99DE-1032130.

XX PR 09-JUL-1999; 99DE-1032186.

XX PR 09-JUL-1999; 99DE-1032206.

XX PR 09-JUL-1999; 99DE-1032227.

XX PR 09-JUL-1999; 99DE-1032228.

XX PR 09-JUL-1999; 99DE-1032229.

XX PR 14-JUL-1999; 99DE-1032230.

XX PR 14-JUL-1999; 99DE-1032922.

XX PR 14-JUL-1999; 99DE-1032926.

```

PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI ) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF71957.
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases.
XX
PS Claim 20; Page 758-759; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 314 AA;

Query Match 64.3%; Score 36; DB 22; Length 314;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10
Db 159 tgdkyfgfl 167
||||:|:

RESULT 6
AAG91391
ID AAG91391 standard; Protein; 334 AA.
XX
AC AAG91391;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 5145.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.

```

```

XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR N-PSDB; AAH66610.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 5145; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 334 AA;

Query Match 64.3%; Score 36; DB 22; Length 334;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGDKFYGLM 10
Db 179 tgdkyfgfl 187
||||:|:

RESULT 7
AAG07312
ID AAG07312 standard; Protein; 349 AA.
XX
AC AAG07312;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4418.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX

```







PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 23-APR-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.

[illegible]

PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149920.  
PR 25-AUG-1999; 99US-0150366.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GDKFYG 8  
|||||  
Db 179 gdkfyg 184

RESULT 11  
AAG07310  
ID AAG07310 standard; Protein; 416 AA.  
XX AC AAG07310;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 4416.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX EPI033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.



PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145182.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 416;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDKFYG 8

Db 197 gdkfyg 202

#### RESULT 12

AAG42889

ID AAG42889 standard; Protein; 416 AA.

XX AC AAG42889;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53542.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-012180.

XX PR 23-MAR-1999; 99US-0123548.

XX PR 25-MAR-1999; 99US-0125788.

XX PR 29-MAR-1999; 99US-0126264.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 16-APR-1999; 99US-0128714.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132486.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 64.3%; Score 36; DB 21; Length 416;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GDKFYG 8  
DB 197 gdkfyg 202

## RESULT 13

AAB29938  
ID AAB29938 standard; Peptide; 15 AA.

XX AAB29938;

DT 09-FEB-2001 (first entry)

DE Human CTLA-4 scaffold protein CDR3 loop (clone 8H).

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;  
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;  
KW diabetic retinopathy; atherosclerosis.

OS Homo sapiens.

PN WO200060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

WPI; 2000-665002/64.

PT Scaffold composed of single-chain polypeptide having beta sandwich  
PT architecture carrying new and randomized peptide sequences useful as  
PT supporting framework and carrying antigen- or receptor binding  
PT fragments -

PS Example 8; Fig 12; 68pp; English.

CC The present invention is concerned with producing scaffold proteins  
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be  
CC used as a scaffold to bind antigen- or receptor-binding fragments. These  
CC can be used in the treatment of diseases such as cancer,  
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and  
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the  
CC production of the proteins of the invention.

SQ Sequence 15 AA;

Query Match 62.5%; Score 35; DB 21; Length 15;

Best Local Similarity 55.6%; Pred. No. 6.5;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DTGDKFYGL 9

DB 6 drgdsyygi 14

## RESULT 14

ABB70933

ID ABB70933 standard; Protein; 274 AA.

XX

AC ABB70933;

XX

DT 26-MAR-2002 (first entry)  
XX Drosophila melanogaster polypeptide SEQ ID NO 39591.  
DE  
XX  
XX Drosophila: developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
KW  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PF  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-0614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX N-PSDB; ABL15036.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PT  
XX  
XX Disclosure: SEQ ID NO 39591; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABE57737-ABE72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 274 AA;  
  
Query Match 62.5%; Score 35; DB 22; Length 274;  
Best Local Similarity 85.7%; Pred. NO. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 GDRFYGL 9  
Db || ||||  
4 gdnfygl 10  
  
RESULT 15  
AAY04108  
ID AAY04108 standard; Protein; 480 AA.  
XX  
XX  
XX AAY04108;  
AC  
XX  
XX 10-JUN-1999 (first entry)  
DT  
XX  
XX Staphylococcus aureus regulator protein.  
DE  
XX  
XX Staphylococcus aureus; regulator; antibacterial; infection; impetigo;  
KW otitis media; cerebral abscess; conjunctivitis; toxic shock syndrome;  
KW wound infection; septic arthritis; transcriptional activator.  
XX  
XX  
OS Staphylococcus aureus.  
XX  
XX  
PN EP893503-A2.  
XX

PD 27-JAN-1999.  
XX  
XX 21-JUL-1998; 98EP-0305801.  
PF  
XX  
XX 25-JUL-1997; 97US-0900953.  
PR  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA  
XX  
XX (SMIK ) SMITHKLINE BEECHAM PLC.  
PA  
XX  
XX Burnham MKR, Lonetto MA, Warren PV;  
PI  
XX  
XX WPI; 1999-097787/09.  
DR  
XX  
XX N-PSDB; AAX19830.  
DR  
XX  
XX New Staphylococcus aureus transcriptional activator (regulator)  
PT polypeptide and polynucleotide - useful as diagnostic reagents and  
PT for prevention and treatment of Staphylococci infections  
PT  
XX  
XX Claim 6; Page 8; 35pp; English.  
PS  
XX  
XX The present sequence represents a transcriptional activator (regulator)  
CC protein isolated from Staphylococcus aureus. Regulator proteins are  
CC useful: (i) for screening for specific inhibitors (antagonists,  
CC potential antibacterial drugs) or activators; (ii) in vaccines to induce  
CC an immunological response (antibody and/or T cell); (iii) to study role  
CC of regulator polynucleotides in pathogenesis; and (iv) to raise anti-  
CC regulator antibodies. Regulator proteins are also useful for treating  
CC conditions that require regulator protein, and antagonists are useful for  
CC treating conditions requiring inhibition of this protein, specifically  
CC infection by S. aureus. Anti-regulator antibodies are used: (i) to  
CC isolate and identify regulator polypeptide-expressing clones; (ii) for  
CC affinity purification of regulator polypeptides; (iii) as therapeutic  
CC antagonists; and (iv) as reagents in diagnostic immunoassays. Regulator  
CC proteins are useful: (i) in genetic immunisation (gene therapy); (ii)  
CC as sources of primers and probes for isolating related sequences, for  
CC diagnosis and staging of infections (particularly by detecting  
CC overexpression of the regulator), for identifying pathogens, and to  
CC detect mutations and polymorphisms (all in standard hybridisation and  
CC amplification tests); (iii) as antisense therapeutics; and (iv) for  
CC identifying epitopes. Diseases diagnosed, prevented or treated include  
CC Staphylococcus aureus infections which cause otitis media; cerebral  
CC abscess; conjunctivitis; toxic shock syndrome; impetigo; wound infection  
CC and septic arthritis.  
XX  
XX  
SQ Sequence 480 AA;  
  
Query Match 62.5%; Score 35; DB 20; Length 480;  
Best Local Similarity 75.0%; Pred. NO. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 TGRFYGL 9  
Db || ||||  
36 tgdwfygm 43  
  
Search completed: June 19, 2002, 08:24:09  
Job time: 246 sec